



Modellversuch

Risikoorientierter Sanktionenvollzug (ROS)

Ergebnisevaluation

Schlussbericht

Christian Schwarzenegger, Patrik Manzoni und Mirjam Baur

Wichtig und für die Kommunikation zu beachten:

Bestimmungen gemäss Zusicherungsverfügung des Bundesamtes für Justiz vom 29. Januar 2010. Auszug: Vor Abschluss der Evaluation ist den Ergebnissen nicht vorzugreifen. Zwischenergebnisse oder Trendmeldungen dürfen nur mit dem Einverständnis des Bundesamtes für Justiz weitergegeben oder veröffentlicht werden. Sämtliche Aktivitäten im Rahmen der Öffentlichkeitsarbeit sind immer vorgängig mit dem BJ abzusprechen. Schlussfolgerungen dürfen vor Genehmigung des Schlussberichts nicht publik gemacht werden.

Kriminologisches Institut der Universität Zürich
Rämistrasse 74/47
CH-8001 Zürich

18. Oktober 2013



Inhaltsverzeichnis

Das Wichtigste in Kürze	IV
Danksagung	V
Abkürzungsverzeichnis	VI
Abbildungsverzeichnis	VII
Tabellenverzeichnis	VIII
1 Einleitung	1
2 Ziele der Ergebnisevaluation	3
2.1 Überprüfung der Güte des Triage-Instruments FaST	3
2.1.1 Objektivität	4
2.1.2 Reliabilität	4
2.1.3 Validität	4
2.2 Vergleich der Rückfälligkeit von prä-ROS-Fällen und ROS-Fällen	4
2.3 Überprüfung der Güte der ROS-Verlaufsliste „Therapie“	5
3 Übersicht der Ergebnisevaluation	6
4 Interraterreliabilität des FaST-Instruments	9
4.1 Methode	9
4.1.1 Stichprobenziehung und Rater	9
4.1.2 Material	10
4.1.3 Durchführung	10
4.1.4 Auswertung	10
4.2 Ergebnisse	10
5 Retrospektive Studie	12
5.1 Methode	12
5.1.1 Informationsquellen	12
5.1.2 Stichprobenziehung	13
5.1.3 Erhebungsinstrumente	15
5.1.4 Beschreibung der erhobenen Variablen	16
5.2 Ergebnisse	21
5.2.1 Beschreibung der Stichprobe	21
5.2.2 Art der Anlassdelikte nach FaST-Kategorien	22
5.2.3 Art der Rückfälligkeit nach FaST-Kategorien	25
5.2.4 Schwere von Gewaltrückfällen nach FaST-Kategorien	28
5.2.5 Mittlere Time-at-risk bis zum Rückfall nach FaST-Kategorien	29
5.2.6 Basisraten der Rückfälligkeit nach FaST-Kategorien und Time-at-risks	30
5.2.7 Time-at-risk-auslösendes Ereignis nach FaST-Kategorien	33
5.2.8 Steuerbares Einkommen nach FaST-Kategorien	36
5.2.9 VRAG nach FaST-Kategorien	42



6	Evaluation von ROS-Verlaufslisten	48
6.1	Ziel und Hintergrund der Verlaufslisten	48
6.2	Entwicklungsphase der Verlaufsliste „Therapie“	49
6.3	Aufbau der ROS-Verlaufsliste „Therapie“	51
6.4	Interraterreliabilität der ROS-Verlaufsliste „Therapie“	52
6.4.1	Methode	52
6.4.2	Ergebnisse IR5	55
6.4.3	Ergebnisse IR6	59
6.5	Validität der ROS-Verlaufslisten „Therapie“ und „Alarm“	66
6.5.1	Methode	66
6.5.2	Ergebnisse	66
7	Konzept für eine prospektive Studie	68
7.1	Stichprobenziehung	69
7.2	Planung und Umsetzung der ROS-Interventionen	70
7.2.1	Planung der Interventionen	70
7.2.2	Umsetzung der Interventionen	70
7.3	Wirkung des ROS-Prozessmodells: Soziale Reintegration	71
7.3.1	Rückfälligkeit	71
7.3.2	Soziales und berufliches Umfeld	72
8	Zusammenfassung und Schlussfolgerungen	73
8.1	Überprüfung der Güte des Triage-Instruments FaST	73
8.2	Erhebung einer Basisrate der Rückfälligkeit von prä-ROS-Fällen	76
8.3	Überprüfung der Güte der ROS-Verlaufsliste „Therapie“	77
8.4	Ausblick auf prospektive Studie	77
8.5	Fazit	79
9	Literatur	80
10	Anhang A: Gewalt- und Sexualdelikte sowie schwerwiegende SVG-Delikte nach FaST	82
11	Anhang B: Beispiel einer FaST-Ausgabe	86
12	Anhang C: Checkpunkte der ROS-Verlaufsliste „Therapie“	87
13	Anhang D: Merkblätter IR5 und IR6	90
14	Anhang E: Detailanalysen auf Item-Ebene, IR5	91
15	Anhang F: Detailanalysen auf Item-Ebene, IR6	106

Das Wichtigste in Kürze

Mit dem Risikoorientierten Sanktionenvollzug (ROS) wird das Ziel verfolgt, den Sanktionenvollzug in ein strukturiertes Prozessmodell zu integrieren, welches sich an empirischen Erkenntnissen orientiert. Im Rahmen des vom Bundesamt für Justiz unterstützten dreijährigen Modellversuchs wurde ROS in der Vollzugspraxis von vier Kantonen (Luzern, St. Gallen, Thurgau und Zürich) erprobt.

Die Evaluation des Modellversuchs beinhaltet eine Prozess- und eine Ergebnisevaluation: Im Fokus der Prozessevaluation steht, wie die Umstrukturierung des Vollzugs in den Modellversuchskantonen von den Mitarbeitenden wahrgenommen und umgesetzt wird. Im Fokus der Ergebnisevaluation stehen die Prüfung der Güte von ROS-Instrumenten und die Bestimmung der Rückfälligkeit vor Einführung von ROS im Hinblick auf eine Einschätzung zur Wirksamkeit des ROS-Regimes. Befriedigende Resultate seitens beider Evaluationszugänge sind zentral für den Einsatz von ROS in der Vollzugspraxis. Im vorliegenden Bericht werden die Ergebnisse der Ergebnisevaluation präsentiert. Die Resultate der Prozessevaluation sind im entsprechenden Bericht von socialdesign enthalten.¹

Das ROS-Instrument Fall-Screening Tool (FaST) wurde von den Bewährungs- und Vollzugsdiensten des Amtes für Justizvollzug Zürich (BVD) entwickelt. Zur Überprüfung der *Güte des FaST* wurden 733 Fälle der Jahre 2004-2005 mit dem FaST gewertet. Anhand derselben Daten wurde eine *Basisrate der Rückfälligkeit von prä-ROS-Fällen* erhoben (via Strafregisterauszüge), die zu einem späteren Zeitpunkt mit der *Rückfälligkeit von ROS-Fällen* verglichen werden soll (Vorher-Nachher-Vergleich). Die Rückfälligkeit von ROS-Fällen kann jedoch erst in mehreren Jahren untersucht werden, da die Sanktion zuerst vollzogen werden muss und danach ein ausreichend langer Untersuchungszeitraum vorliegen muss, welcher eine Beurteilung über die Rückfälligkeit zulässt.

Schliesslich wurde das gemeinsam von den BVD und des Psychiatrisch-Psychologischen Diensts des Amtes für Justizvollzug Zürich (PPD) entwickelte ROS-Instrument „Verlaufsliste Therapie bzw. Alarm“ in zwei Studien anhand von 30 Fällen auf Verlässlichkeit und Gültigkeit überprüft.

Die Erhebung der Daten für die Ergebnisevaluation war teilweise herausfordernd, was insbesondere auf die heterogene Dokumentationsweise (die relevanten Informationen mussten aus mehreren Quellen zusammengetragen werden) und den heterogenen Grad an Dokumentation der Fälle (nicht alle Fälle waren gleich ausführlich dokumentiert) zurückzuführen war. An einigen Stellen ist somit die Repräsentativität der Stichprobe eingeschränkt, worauf an den gegebenen Stellen hingewiesen wird.

Insgesamt hat die Ergebnisevaluation bezüglich der Güte der ROS-Instrumente vielversprechende Resultate ergeben. Noch offen ist, wie wirksam ROS bezüglich der Verhinderung von Rückfällen ist. Die Rückfälligkeit von ROS-Fällen sollte im Rahmen einer prospektiven Studie erhoben werden. Erst wenn die Rückfälligkeitsrate der ROS-Fälle eruiert werden kann, ist eine Aussage zur Wirksamkeit von ROS bezüglich der Verhinderung von Rückfällen möglich.

¹ Die Prozessevaluation wurde unabhängig von uns durch die Firma socialdesign durchgeführt.

Danksagung

An dieser Stelle bedanken wir uns bei allen in den Modellversuch involvierten Personen, die uns in verschiedenen Phasen der Ergebnisevaluation unterstützt haben. Bei Alex Schilling, Daniel Treuthardt, Klaus Mayer, Hans-Jürg Patzen, Jérôme Endrass und Astrid Rossegger bedanken wir uns insbesondere für die konstruktiven Hinweise und Diskussionen bezüglich der Planung und Umsetzung einzelner Arbeitsschritte. Die Datenerhebung wurde uns durch die fachlichen Kenntnisse von Gregor Wanner und Daniel Zobrist sehr erleichtert, wofür wir uns bedanken möchten. Nicht zuletzt danken wir Damaris Reichen und Erika Wittwer für ihre koordinative und administrative Unterstützung.



Abkürzungsverzeichnis

ANAG	Gesetz über den Aufenthalt und die Niederlassung von Ausländern (alt)
AUC	Area under the curve
AuG	Ausländergesetz
BetmG	Betäubungsmittelgesetz
BF	Standardisiertes Therapieberichtsformular
BJ	Bundesamt für Justiz
BVD	Bewährungs- und Vollzugsdienste des Amtes für Justizvollzug Zürich
CATS	Childhood and Adolescent Taxon Scale
CR!MS	Computergestütztes Risiko-Management-System
FaST	Fall-Screening Tool
FOTRES	Forensisch Operationalisiertes Risiko-Therapie-Evaluations-System
G/S-Delikte	Gewalt- oder Sexualdelikte
IR1-IR6	Interraterreliabilitätsstudien 1-6
JuV	Amt für Justizvollzug des Kantons Zürich
KI	(asymptotisches) Konfidenzintervall
LERF	Labile eigenständig risikorelevante Faktoren
LSI-R	Level of Service Inventory Revised
MSTG	Militärstrafgesetz (AMSTG: Altes Militärstrafgesetz)
PCL-R	Psychopathy-Checkliste Revised
PIC-R	Personal, Interpersonal and Community-Reinforcement Approach
PPD	Psychiatrisch-Psychologischer Dienst des Amtes für Justizvollzug Zürich
RIS	Rechtsinformationssystem der Direktion der Justiz und des Innern des Kanton Zürich)
ROC	Receiver operating characteristics
ROS	Risikoorientierter Sanktionenvollzug
StGB	Schweizerisches Strafgesetzbuch (A StGB: Altes Schweizerisches Strafgesetzbuch)
StPO	Schweizerische Strafprozessordnung
SVG	Strassenverkehrsgesetz
TSchG	Tierschutzgesetz
VL	ROS-Verlaufslisten
VOSTRA	Vollautomatisiertes Strafregister
VRAG	Violence Risk Appraisal Guide



Abbildungsverzeichnis

Abbildung 1: ROS-Prozessmodell. Aus: Amt für Justizvollzug Kanton Zürich (2012)	1
Abbildung 2: Übersicht Hauptgütekriterien	3
Abbildung 3: Stichprobe nach Alter	21
Abbildung 4: Mittleres Einkommen (Mediane) in CHF, nach FaST-Kategorien und Time-at-risks	40
Abbildung 5: Verteilung der VRAG-Risikoklassen.....	44
Abbildung 6: ROS-Prozessmodell. Aus: Amt für Justizvollzug Kanton Zürich (2012)	48
Abbildung 7: Planung, Umsetzung und Wirkung von ROS-Interventionen.....	68
Abbildung 8: Merkblatt für die Rater, exemplarisch für IR6	90

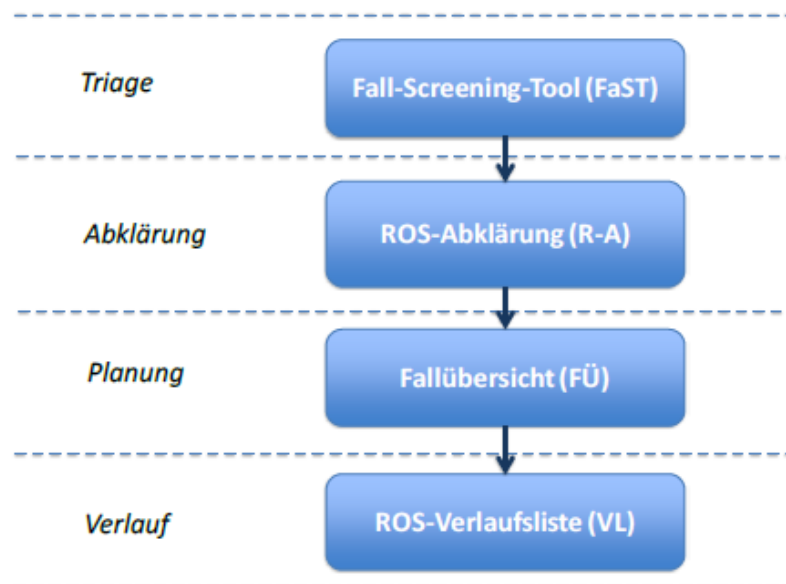
Tabellenverzeichnis

Tabelle 1: Überblick Ergebnisevaluation	6
Tabelle 2: Ziele der Ergebnisevaluation und Studien, um diese zu erreichen	8
Tabelle 3: Interraterreliabilität FaST	10
Tabelle 4: Grundpopulation und Stichprobe der retrospektiven Studie.....	14
Tabelle 5: Ausschlussgründe	14
Tabelle 6: VRAG-Risikoklassen und Rückfallwahrscheinlichkeiten	16
Tabelle 7: Art des Anlassdelikts, differenziert nach FaST-Kategorien.....	22
Tabelle 8: G1-Anlassdelikte, differenziert nach FaST-Kategorien	23
Tabelle 9: G2-Anlassdelikte, differenziert nach FaST-Kategorien	24
Tabelle 10: G3-Anlassdelikte, differenziert nach FaST-Kategorien	24
Tabelle 11: Kombinationsmöglichkeiten von Rückfallarten und FaST-Kategorien	25
Tabelle 12: Rückfallart, differenziert nach FaST-Kategorien.....	26
Tabelle 13: Vorliegen eines G/S-Delikts (Vorstrafe oder Anlassdelikt), differenziert nach FaST-Kategorien	27
Tabelle 14: Schwere von Gewaltrückfällen, differenziert nach FaST-Kategorien	28
Tabelle 15: Mittlere Time-at-risks in Monaten bis zur Rückfälligkeit, differenziert nach FaST-Kategorien	29
Tabelle 16: Rückfälligkeit nach verschiedenen Time-at-risks, differenziert nach FaST-Kategorien.....	30
Tabelle 17: Rückfälligkeit, differenziert nach FaST-Kategorien, 1 Jahr Time-at-risk	31
Tabelle 18: Rückfälligkeit, differenziert nach FaST-Kategorien, 3 Jahre Time-at-risk	31
Tabelle 19: Rückfälligkeit, differenziert nach FaST-Kategorien, 6 Jahre Time-at-risk	32
Tabelle 20: Ereignisse, welche die Time-at-risk auslösten, differenziert nach FaST-Kategorien.....	34
Tabelle 21: Art der Entlassung aus dem Strafvollzug, differenziert nach FaST-Kategorien	36
Tabelle 22: Ausschlüsse und Stichprobengrößen für Auswertungen zu Steuerdaten	37
Tabelle 23: Einkommen nach verschiedenen Time-at-risks, differenziert nach FaST-Kategorien.....	38
Tabelle 24: Mittleres Einkommen in CHF, differenziert nach FaST-Kategorien und Time-at-risks.....	40
Tabelle 25: Anwendung des VRAG, differenziert nach FaST-Kategorie	43
Tabelle 26: Verteilung der VRAG-Risikoklassen	44
Tabelle 27: Statistische Verteilung der VRAG-Werte, differenziert nach FaST-Kategorien	45
Tabelle 28: VRAG-Risikoklassen und Anlassdelikte, differenziert nach FaST-Kategorien	46
Tabelle 29: Formen der Risiko-Abklärung	47
Tabelle 30: Beispiele zur Ausprägung des Kappa-Koeffizienten, abhängig von der Prävalenz.....	54
Tabelle 31: Items der VL „Therapie“ und die zugehörigen Kappa-Koeffizienten, IR5.....	55
Tabelle 32: Zusammenfassung der Übereinstimmung, IR1 bis IR5	56
Tabelle 33: Erkennung von Hinweisen auf einen kritischen Therapieverlauf, IR5.....	58
Tabelle 34: Items der VL „Therapie“ und die zugehörigen Kappa-Koeffizienten, IR6.....	59
Tabelle 35: Zusammenfassung der Übereinstimmung, IR1 bis IR6	60
Tabelle 36: Erkennung von Hinweisen auf einen kritischen Therapieverlauf, IR6.....	65
Tabelle 37: Interraterreliabilität und konkurrente Validität, Items 20-29	67
Tabelle 38: Interraterreliabilität und konkurrente Validität, Items 24-29	67
Tabelle 39: Gewalt-, Sexualdelikte sowie schwere SVG-Delikte nach FaST-Manual.....	82

1 Einleitung

Die zentrale Herausforderung bei der Planung und Koordination des Sanktionenvollzugs besteht darin, die relevanten Informationen von allen involvierten Akteuren einzubeziehen, zumal es sich um ein interdisziplinäres und multiinstitutionelles Feld handelt. Der Risikoorientierte Sanktionenvollzug (ROS) verfolgt das Ziel, bisher vorhandene Schnittstellenproblematiken zu reduzieren, indem die einzelnen Arbeitsschritte der Vollzugsplanung klar definiert werden und in den Kontext eines mehrstufigen **Prozessmodells** gestellt werden. Damit soll dem zentralen Ziel von ROS, eine rückfallpräventive und sozial reintegrierende Wirkung zu erzielen, Rechnung getragen werden (siehe zum Ganzen Amt für Justizvollzug Kanton Zürich, 2011). Abbildung 1 stellt das ROS-Prozessmodell dar.

Abbildung 1: ROS-Prozessmodell. Aus: Amt für Justizvollzug Kanton Zürich (2012)



Zu Beginn des Prozesses erfolgt eine Triage sämtlicher Fälle durch das eigens für ROS konzipierte Instrument, das Fall-Screening Tool (FaST). Dieses gibt die drei Stufen A, B oder C aus, welche folgendermassen interpretiert werden sollen (Bewährungs- und Vollzugsdienste Zürich, Fachsupport & Lernprogramme, 2013):

- **A-Fälle:** Geringer Abklärungsbedarf
- **B-Fälle:** Erhöhter Abklärungsbedarf bezüglich allgemeiner Delinquenz
- **C-Fälle:** Erhöhter Abklärungsbedarf bezüglich Gewalt- und Sexualdelinquenz

Es ist wichtig festzuhalten, dass das FaST kein eigentliches Risikoproggnoseinstrument darstellt, sondern ein mit relativ geringem Aufwand auszufüllendes Instrument sein soll, das jene Fälle anzeigt, die einen besonderen Abklärungsbedarf aufweisen. Da aber das FaST so konzipiert ist, dass viele rückfallrelevante Faktoren erfragt werden, könnte man von einem indirekten bzw. impliziten Rückfallprognoseinstrument sprechen. Die gängige Modellversuchspraxis besteht darin, dass lediglich C-Fälle, welche eine Freiheitsstrafe von mehr als sechs Monaten zu verbüssen haben, standardmässig die

weiteren Stufen des Prozessmodells erreichen (Bewährungs- und Vollzugsdienste Zürich, Projekt- und Qualitätsmanagement, 2012).

Für diese erfolgt im Rahmen der ROS-Abklärung ein kombiniertes Risiko- und Bedarfsassessment durch die „Abteilung für Forensisch-Psychologische Abklärungen“ der Bewährungs- und Vollzugsdienste des Amts für Justizvollzug Zürich (BVD). Die Intervention selbst wird dann durch den Fallverantwortlichen koordiniert. Hierzu führt er die ROS-Abklärung und andere gesammelte Informationen (z.B. relevante Informationen von Vollzugsbehörden) in der sog. Fallübersicht zusammen. Der Verlauf des Sanktionenvollzugs wird sodann standardisiert, mittels sog. ROS-Verlaufslisten (VL), überprüft.

Das Kriminologische Institut der Universität Zürich hat den Auftrag, die Ergebnisevaluation des vom Bund geförderten Modellversuchs ROS durchzuführen, womit zu den Zielen der Ergebnisevaluation übergeleitet werden soll.

2 Ziele der Ergebnisevaluation

Die Ergebnisevaluation verfolgt drei Hauptziele, welche nachfolgend einzeln erörtert werden.

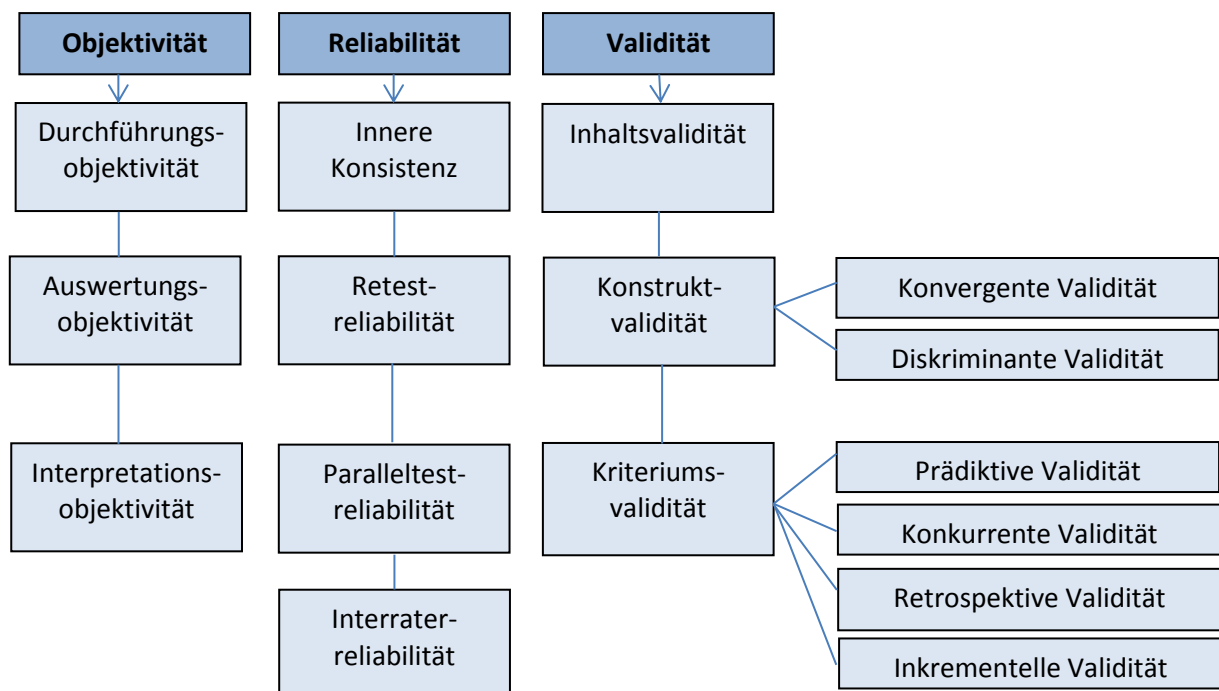
- Überprüfung der Güte des Triage-Instruments FaST
- Vergleich der Rückfälligkeit von prä-ROS-Fällen und ROS-Fällen
- Überprüfung der Güte der ROS-Verlaufsliste „Therapie“

2.1 Überprüfung der Güte des Triage-Instruments FaST

Wie aus dem ROS-Prozessmodell ersichtlich wird, kommt dem Triage-Instrument FaST eine zentrale Bedeutung zu: Es bestimmt darüber, ob weitere Abklärungen vorgenommen werden oder nicht. Es ist somit wichtig, dass das FaST Fälle mit einem erhöhten Abklärungsbedarf bezüglich Gewalt- und Sexualdelinquenz tatsächlich als „C“ identifiziert.

Die Güte eines Instruments wird anhand der sog. Hauptgütekriterien bestimmt. Dabei unterscheidet man die Objektivität, Reliabilität und Validität, welche alle drei weiter differenzierbar sind. Abbildung 2 gibt eine Übersicht über die Hauptgütekriterien samt Differenzierungen. Es werden jedoch nachfolgend lediglich die drei Hauptgütekriterien, ohne Differenzierungen, erläutert. Die für die Ergebnisevaluation relevanten Differenzierungen werden an jenen Stellen im Bericht erörtert, wo diese zur Anwendung kommen.

Abbildung 2: Übersicht Hauptgütekriterien



Zusammenstellung nach Bühner (2011); Hussy, Schreier und Echterhoff (2013).

2.1.1 Objektivität

Die Objektivität drückt aus, inwiefern die Ergebnisse eines Tests unabhängig vom Untersucher sind (Bühner, 2011). Diese Unabhängigkeit sollte sowohl die Durchführung, die Auswertung als auch die Interpretation eines Tests betreffen. Die Erfassung der Objektivität ist weniger Teil der Ergebnisevaluation, da diese eher eingeschätzt als gemessen werden kann.

An dieser Stelle soll jedoch in aller Kürze darauf hingewiesen sein, dass die Objektivität des FaST durch die Existenz eines Manuals (Durchführungsobjektivität), die automatisierte Auswertung (Auswertungsobjektivität) und die definierte Regelung, wie bei welcher FaST-Kategorie zu reagieren ist (Interpretationsobjektivität), gut gestützt ist.

2.1.2 Reliabilität

Die Reliabilität gibt den Grad der Genauigkeit an, mit dem ein Test ein Merkmal misst (Bühner, 2011). Dabei ist es unwichtig, ob es sich bei diesem Merkmal auch tatsächlich um jenes handelt, welches ein Test zu messen beansprucht. Die Reliabilität kann auf verschiedene Arten erfasst werden. Im Rahmen der Ergebnisevaluation wird die Erfassung der Interraterreliabilität verwendet, welche an gegebener Stelle erläutert wird.

2.1.3 Validität

Die Validität zeigt auf, ob ein Test das misst, was er zu messen vorgibt. Eine zufriedenstellende Reliabilität ist Voraussetzung für das Vorliegen von Validität: Misst ein Test ein gewisses Merkmal genau (zufriedenstellende Reliabilität), so kann mit der Erhebung der Validität überprüft werden, ob dieses Merkmal auch das misst, was der Test zu messen beabsichtigt. Wie aus Abbildung 2 hervorgeht, gibt es drei Differenzierungen der Validität (Inhaltsvalidität, Konstruktvalidität, Kriteriumsvalidität), die ihrerseits nochmals unterteilt werden können. Für die Ergebnisevaluation des FaST werden die prädiktive und die konvergente Validität überprüft, welche anhand der Daten an gegebener Stelle erläutert werden.

2.2 Vergleich der Rückfälligkeit von prä-ROS-Fällen und ROS-Fällen

Nebst der Überprüfung der Güte des FaST verfolgt die Ergebnisevaluation das Ziel, die Rückfälligkeit von prä-ROS- und ROS-Fällen zu vergleichen. Die Rückfälligkeit fungiert als Hauptindikator für die soziale Reintegration.²

Eine Basisrate der Rückfälligkeit von prä-ROS-Fällen wurde anhand von Fällen der Jahre 2004/2005 erhoben (nähere Angaben hierzu siehe Kapitel 5). Dabei wurden Time-at-risks (Zeitspanne, in der eine Person die Gelegenheit hatte, rückfällig zu werden) von einem, drei und sechs Jahren untersucht. Diese Basisrate soll mit der Rückfälligkeit von ROS-Fällen verglichen werden, welche im Rahmen einer prospektiven Studie erhoben werden soll. Dies wird erst in mehreren Jahren möglich sein. Der vorliegende Bericht beinhaltet eine Planung einer prospektiven Studie (Kapitel 7).

² Als weiterer Indikator wird das steuerbare Einkommen erhoben.

2.3 Überprüfung der Güte der ROS-Verlaufsliste „Therapie“

Ein weiteres Ziel der Ergebnisevaluation besteht in der Bestimmung der Güte der VL „Therapie“. Wie beim FaST stehen die Überprüfung der Reliabilität (anhand der Interraterreliabilität) und Validität (konkurrente Validität) im Vordergrund.

Nachfolgend wird eine Übersicht über die verschiedenen Etappen der Ergebnisevaluation gegeben, welche sodann mit den drei Hauptzielen in Verbindung gebracht werden.

3 Übersicht der Ergebnisevaluation

Am 1. Mai 2010 startete der Modellversuch „Risikoorientierter Sanktionenvollzug“. Ein Jahr später, am 1. Mai 2011 begann das Kriminologische Institut mit der Ergebnisevaluation. Diese bestand aus vier Komponenten, welche der nachfolgenden Darstellung (Tabelle 1) zu entnehmen sind. Im Rahmen von zwei Zwischenberichten wurden die bis zu den damaligen Zeitpunkten vorliegenden Resultate berichtet (Schwarzenegger, Manzoni & Baur, 2012; Baur, Manzoni & Schwarzenegger, 2013). Der vorliegende Schlussbericht stellt eine umfassende Darstellung der Ergebnisse der Evaluation dar.

Tabelle 1: Überblick Ergebnisevaluation

2011			2012				2013		
Q 2	Q 3	Q 4	Q 1	Q 2	Q 3	Q 4	Q 1	Q 2	Q 3
1. Vorbereitung der Evaluation und Interraterreliabilitätsstudie FaST*									
		2. Retrospektive Studie* (Validität FaST und Basisrate Rückfälligkeit)							
							3. Planung der prospektiven Studie*		
							4. Evaluation von ROS-Verlaufslisten*		

Q = Quartal

*Zwischenberichte und Schlussbericht:

01. 03. 2012: Zwischenbericht I

29. 04. 2013: Zwischenbericht II

01. 10. 2013: Schlussbericht

Wie Tabelle 1 zeigt, wurde das FaST zuerst im Rahmen einer **Interraterreliabilitätsstudie** auf seine Genauigkeit geprüft (siehe Kapitel 4). Dabei wurde überprüft, inwiefern unterschiedliche Anwender (sog. Rater) des FaST bei Bewertung derselben Fälle in den Resultaten (A, B oder C) miteinander übereinstimmen. Die Reliabilität ist, wie erwähnt, ein Hauptgütekriterium. Die Interraterreliabilitätsstudie dient demnach dem Ziel, die Güte des FaST zu bestimmen (siehe Tabelle 2). Weiter standen in dieser Phase Vorbereitungen für die retrospektive Studie an (z.B. Einholen von Bewilligungen für Zugänge zu Datenbanken).

In einem zweiten Schritt wurden Daten von 733 Personen erhoben, deren Sanktionenvollzug durch die BVD durchgeführt und in den Jahren 2004 oder 2005 abgeschlossen wurde (siehe hierzu genauere Ausführungen in Kapitel 5.1.2). Das FaST wurde retrospektiv für diese Personen angewandt, das heisst ausgegangen wurde von den Informationen, welche bereits in den Jahren 2004 bzw. 2005 vorlagen. In Anlehnung an den Erhebungszeitpunkt nennen wir diese Studie **retrospektive Studie**. Sie dient der Überprüfung der Validität des FaST und somit dem Ziel, die Güte des FaST zu überprüfen. Zentral ist der Zusammenhang der FaST-Kategorien und der Rückfälligkeit (als Hauptindikator für die soziale Reintegration) der 733 Personen. Als weiterer Indikator der sozialen Reintegration dient das erzielte steuerbare Einkommen nach Verbüssung der Sanktion. Hierbei handelt es sich um die Überprüfung der sog. prädiktiven Validität: Es wird ermittelt, inwiefern Zusammenhänge eines Testergebnisses (FaST) mit einem später ermittelten Kriterium (Rückfall und erzielt steuerbares Einkommen) in Zusammenhang stehen.

Als zweite Form der Validitätsprüfung wurde die konvergente Validität ermittelt. Dabei wird geprüft, wie ein neues Instrument (FaST) mit einem etablierten Instrument von gleichem Gültigkeitsbereich in Zusammenhang steht. Als solches eignete sich der „Violence Risk Appraisal Guide“ (VRAG; Quinsey, Harris, Rice & Cormier, 2006). Weitere Ausführungen zum VRAG finden sich in Kapitel 5.1.3.2.

Die in der retrospektiven Studie ermittelten Rückfälligkeitsraten dienen zudem dem zweiten Ziel der Ergebnisevaluation, dem Vergleich der Rückfälligkeit von prä-ROS-Fällen und ROS-Fällen. Die ermittelte Basisrate der Rückfälligkeit von prä-ROS-Fällen wird in mehreren Jahren mit der Rückfälligkeit von ROS-Fällen verglichen (prospektive Studie). Dabei interessieren insbesondere C-Fälle, da vorwiegend bei diesen gemäss ROS-Prozessmodell interveniert wird. Die retrospektive Studie wird in Kapitel 5 thematisiert.

Neben der Erhebung der Rückfälligkeit von ROS-Fällen, welche mit der in der retrospektiven Studie ermittelten Basisrate verglichen werden, schlägt das Konzept einer **prospektiven Studie** ebenfalls eine Erfassung der Planung und Umsetzung von Interventionen vor. Eine regelmässige wissenschaftliche Überprüfung der ROS-Prozesse ist unumgänglich, wenn die Qualitätssicherung in der Vollzugspraxis gewährleistet werden soll.

Nicht von Beginn an geplant war die **Evaluation der VL „Therapie“**. Diese wird bei Gewalt- und Sexualstraftätern angewandt, die eine ambulante oder stationäre Therapie absolvieren. Ziel ist es, Schnittstellenproblematiken zwischen den Vollzugsdiensten und Therapeuten zu minimieren. Konkret helfen sie der fallverantwortlichen Person der Vollzugsdienste, die regelmässig vom Therapeuten gelieferten Therapieberichte auf ihre Vollständigkeit und inhaltliche Kohärenz zu prüfen sowie Hinweise auf einen kritischen Therapieverlauf zu erkennen.

Die Evaluation der VL beinhaltet wiederum die Messung der Interraterreliabilität und die Überprüfung der Validität in Form der konkurrenten Validität. Letztere macht eine Aussage dazu, inwiefern zwei Kriterien, welche in einem ähnlichen Zeitraum erhoben wurden, miteinander übereinstimmen. Die Evaluation der VL „Therapie“ wird in Kapitel 5.2 thematisiert. Nachfolgende Darstellung zeigt auf, mit welcher Studie welche Ziele verfolgt werden.



Tabelle 2: Ziele der Ergebnisevaluation und Studien, um diese zu erreichen

Ziel der Ergebnisevaluation	Spezifizierung	Studie, um Ziel zu erreichen
1. Überprüfung der Güte des Triage-Instrumentes FaST	<ul style="list-style-type: none">- Interraterreliabilität- Prädiktive Validität- Konvergente Validität	<ul style="list-style-type: none">→ Interraterreliabilitätsstudie FaST→ Retrospektive Studie→ Retrospektive Studie
2. Vergleich der Rückfälligkeit von prä-ROS-Fällen und ROS-Fällen		<ul style="list-style-type: none">→ Retrospektive Studie→ Planung der prospektiven Studie
3. Überprüfung der Güte von Verlaufslisten (Reliabilität, Validität)	<ul style="list-style-type: none">- Interraterreliabilität- Konkurrente Validität	<ul style="list-style-type: none">→ Evaluation von ROS-Verlaufslisten→ Evaluation von ROS-Verlaufslisten

4 Interraterreliabilität des FaST-Instruments

Die Interraterreliabilität als Form der Reliabilität macht wie erwähnt eine Aussage darüber, inwiefern verschiedene Anwender eines Instrumentes (sog. Rater) auf dieselben Resultate kommen, wenn sie dieselben Daten mit demselben Instrument codieren. Es geht also um die Aussage, wie stabil oder konsistent Messungen zwischen mehreren Ratern sind. Eine zufriedenstellende Interraterreliabilität ist Voraussetzung für den praktischen Einsatz eines Instrumentes. Eine unzureichende Interraterreliabilität führt dazu, dass Unterschiede in den Messungen nicht eindeutig den Versuchsobjekten (hier: den Fällen) zugeordnet werden können. Vielmehr ist ein Teil der Unterschiede dadurch zustande gekommen, dass das Instrument keine eindeutigen Zuordnungen zulässt und somit bei verschiedenen Ratern andere Ergebnisse resultieren. Die Interraterreliabilität fungiert also als Qualitätsindikator der Reproduzierbarkeit von Messungen.

4.1 Methode

4.1.1 Stichprobenziehung und Rater

Im Oktober 2011 erfolgte eine Stichprobenziehung von 40 Fällen, die bereits von einem BVD-Mitarbeitenden mittels dem FaST codiert wurden. Dazu wurde zufällig ein Mitarbeitender der Abteilung Strafvollzug gewählt.³ Die Abteilung Strafvollzug regelt und koordiniert den Vollzug von Freiheitsstrafen in Gefängnissen und Strafanstalten. Sie weist zwei Charakteristika auf, welche für eine Zufallsauswahl von Vorteil sind: Zum einen wird in dieser Abteilung eine heterogene Gruppe von Straftätern betreut. Dies hat den Vorteil, dass davon ausgegangen werden kann, dass der ermittelte Wert der Interraterreliabilität nicht nur für eine Subpopulation von Straftätern gilt. Zum anderen hatten zum Zeitpunkt der Stichprobenziehung alle Mitarbeitenden dieser Abteilung bereits mehr als 40 Fälle mit dem FaST codiert, was die Zufallsauswahl eines Mitarbeitenden erleichterte. Im nächsten Schritt wurde die Anzahl Fälle, welche der betreffende Mitarbeitende seit Beginn von ROS mit dem FaST codiert hatte, durch 40 geteilt. Der daraus resultierende Quotient gab Aufschluss darüber, jeder wievielte codierte Fall in die Stichprobe aufzunehmen war.

Die 40 Fälle wurden zusätzlich durch zwei Mitarbeitende des Kriminologischen Instituts gewertet, so dass die Berechnung der Interraterreliabilität basierend auf den Wertungen von drei Ratern vorgenommen werden konnte.

³ Neben der Abteilung Strafvollzug gibt es die Abteilungen gemeinnützige Arbeit, Schuldensanierung, ambulante und stationäre Behandlungen bei Alkoholabhängigkeit, Betäubungsmittelabhängigkeit und psychischen Störungen, den Straf- und Massnahmenvollzug 3, welcher aus einem spezialisierten Team für Straffällige mit hohem Gefährlichkeits- und Rückfallpotential besteht, sowie die Abteilung des Straf- und Massnahmenvollzugs für junge Erwachsene. Eine Übersicht findet sich auf:

http://www.justizvollzug.zh.ch/internet/justiz_innere/juv/de/ueber_uns/organisation/bvd.contenttab.2.bereiche.html#a-content-content-internet-justiz_innere-juv-de-ueber_uns-organisation-bvd-jcr-content-contentPar-contenttab, besucht am 9.8.2013.

4.1.2 Material

Die für die FaST-Codierung relevanten Urteile und Strafregisterauszüge wurden für die beiden Rater des Kriminologischen Instituts zusammengestellt und zur Verfügung gestellt. Zur Anwendung des FaST werden zudem Angaben aus einer kantonalen Datenbank, dem „Rechtsinformationssystem der Direktion der Justiz und des Innern des Kantons Zürich“ (RIS) benötigt. Diese konnten durch die Evaluatoren selbst gesucht und ausgedruckt werden. Welche Informationen für das FaST benötigt werden, wird erst in Kapitel 5.1.4.1 thematisiert, da die Inhalte des FaST für die Überprüfung der Interraterreliabilität nicht im Zentrum stehen.

4.1.3 Durchführung

Die beiden Mitarbeitenden des Kriminologischen Instituts führten das FaST bei den 40 Fällen unabhängig voneinander durch. Um gleiche Voraussetzungen zu schaffen, erhielten beide Mitarbeitenden vorab das FaST-Manual sowie die gleiche Schulung wie die BVD-Mitarbeitenden vor ihrem ersten FaST-Einsatz. In der Folge codierten die beiden Mitarbeitenden die 40 Fälle. Die resultierenden FaST-Ausprägungen (A, B oder C) wurden auf Übereinstimmung zwischen den drei Ratern überprüft.

4.1.4 Auswertung

Zur Bestimmung der Interraterreliabilität existiert eine ganze Reihe verschiedener Reliabilitätskoeffizienten. Im Rahmen der vorliegenden Studie wurde Krippendorff's α gewählt, welches im Vergleich mit anderen Reliabilitätskoeffizienten verschiedene Vorteile aufweist. So ist es unabhängig von der Anzahl Rater sowie vom Skalenniveau der erhobenen Daten und kann auch eingesetzt werden, wenn einzelne Daten fehlen. Weiter nimmt es eine Korrektur von Zufallsergebnissen vor (Hayes & Krippendorff, 2007). Krippendorff's α kann einen Wert zwischen 0 und 1 annehmen. Werte ab 0.8 gelten als sehr gute Übereinstimmung. Werte zwischen 0.6 und 0.8 sind als akzeptable bis gute Übereinstimmung zu beurteilen und Werte, die kleiner sind als 0.6, gelten als inakzeptabel (Krippendorff, 1980).

4.2 Ergebnisse

Alle drei Rater codierten alle 40 Fälle, so dass es keine fehlenden Daten gab. Neben der wohl wichtigsten Überprüfung der Interraterreliabilität über alle drei Rater wurden jeweils ebenfalls die Werte über je zwei der Rater bestimmt. Tabelle 3 gibt einen Überblick über die Resultate.

Tabelle 3: Interraterreliabilität FaST

Kombination Rater	Krippendorff's Alpha
Alle 3 Rater	.84
BVD-MA und Krim. 1	.80
BVD-MA und Krim. 2	.87
Krim 1 und Krim. 2	.84

Anmerkung: N = 40. BVD-MA = BVD-Mitarbeitender, Krim. 1 und Krim 2 = Mitarbeitende des Kriminologischen Instituts.

Alle Reliabilitätskoeffizienten zeigen eine sehr gute Übereinstimmung. Dies deutet darauf hin, dass die Kodierungsanweisungen im FaST-Manual gut dokumentiert sind, so dass eine hohe Übereinstimmung zwischen verschiedenen Anwendern erzielt wird.

Bei sechs der 40 Fälle wich jeweils ein Rater von den anderen beiden Ratern ab. Die betreffenden Fälle wurden nochmals von den beiden Mitarbeitenden des Kriminologischen Instituts geprüft und – wo Klärungsbedarf bestand – zusammen mit einem Mitarbeitenden, welcher massgeblich an der Entwicklung des FaST beteiligt war, besprochen, um die Ursachen für die Diskrepanzen zu eruieren. Diese konnten in allen Fällen eindeutig aufgeklärt werden:

- *Vier Fälle* wurden vom BVD-Mitarbeitenden mit einer älteren CR!MS-Version⁴ bearbeitet, da die Fälle schon einige Zeit zurück lagen. Einzelne Items wurden in der älteren CR!MS-Version stärker oder schwächer gewichtet, so dass dies im Resultat zu einer anderen Ausprägung führte.
- Bei *einem Fall* wurde von zwei Ratern ein Tatbestand (Drohung, Art. 180 StGB) in einem Urteil übersehen, was den Ausschlag für eine andere Ausprägung gab.
- *Ein Fall* hatte neben einem Alias-Namen ebenfalls ein Alias-Geburtsdatum. Richtete man sich nach dem „echten“ Geburtsdatum, so ist ein jugendanwaltschaftlicher Eintrag fälschlicherweise entstanden (Person war eigentlich zu alt, um nach Jugendstrafgesetz verurteilt zu werden). Einer der Rater hat sich an dem „echten“ Geburtsdatum orientiert, die anderen beiden am Alias-Geburtsdatum. Da die Anzahl Einträge ins Jugendstrafregister relevant für die Gewichtung des FaST ist, führte diese Diskrepanz zu unterschiedlichen FaST-Ausprägungen.

Zusammenfassend kann gesagt werden, dass das FaST mit Hilfe einer Schulung und des Manuals so durchgeführt werden kann, dass es bei verschiedenen Ratern zu den gleichen Ergebnissen führt. Somit ist eine zentrale Voraussetzung für die Anwendung in der Praxis gegeben. In einem nächsten Schritt kann die Validität des FaST untersucht werden.

⁴ Das „Computergestütztes Risiko-Management-System“ (CR!MS) ist ein Excel-basiertes Tool, welches neben dem FaST auch andere ROS-Instrumente enthält.

5 Retrospektive Studie

Zur Bestimmung der Validität des FaST interessiert insbesondere, ob das FaST jene Personen als C-Fälle einstuft, die aufgrund ihrer späteren Rückfälligkeit einen besonderen Abklärungsbedarf nötig gehabt hätten.

Im Rahmen der retrospektiven Studie wurde anhand abgeschlossener Fälle einerseits die prädiktive, andererseits die konvergente Validität untersucht. Zur Ermittlung der prädiktiven Validität wurden die Rückfallraten und das steuerbare Einkommen nach Verbüßung der Sanktion als Indikatoren für die soziale Integration untersucht. Dabei wurde nach FaST-Kategorie differenziert. Mit anderen Worten wurde überprüft, inwiefern das FaST die soziale Reintegration vorhersagen konnte. Von besonderem Interesse war, inwiefern C-Fälle gehäuft mit einem Gewalt- oder Sexualdelikt (G/S-Delikt) rückfällig wurden, wobei Gewaltrückfälle zudem nach ihrer Schwere unterteilt wurden. Die konvergente Validität wurde durch die Ermittlung des Zusammenhangs zwischen FaST und VRAG untersucht.

Weiter wurde anhand dieser Fälle die Basisrate der Rückfälligkeit vor Implementierung von ROS ermittelt (prä-ROS-Fälle). Es wurde zudem erhoben, welche Art von Rückfall (G/S-Delinquenz oder allgemeine Delinquenz) vorlag sowie nach welchem Zeitraum und wie häufig die Person rückfällig wurde. Diese Basisrate der Rückfälligkeit soll in der prospektiven Studie mit der Rückfälligkeit aktueller Fälle (ROS-Fälle) verglichen werden, um so einen Hinweis auf eine verbesserte Rückfallprävention und somit die Wirksamkeit des ROS-Prozessmodells zu erhalten.

5.1 Methode

5.1.1 Informationsquellen

Die BVD ermöglichten dem Kriminologischen Institut den Zugang zu von ihnen administrierten Fällen. Hierzu unterzeichneten die Mitarbeitenden des Kriminologischen Instituts Geheimhaltungsverpflichtungen. Es waren zwei Datenbanken von zentraler Bedeutung. Das bereits erwähnte RIS ermöglichte es, sowohl die Stichprobenziehung (siehe nachfolgendes Kapitel 5.1.2) als auch die Erhebung delikt- und vollzugsrelevanter Informationen durchzuführen. Da das RIS eine kantonale Datenbank darstellt, war es für die Erhebung der Rückfälligkeit nötig, eine bundesweite Informationsquelle heranzuziehen. Als solche bot sich das „vollautomatisierte Strafregister“ (VOSTRA) an. Daraus waren Informationen zum Datum der Verurteilung, dem Straftatbestand, der Zeitspanne der Deliktbegehung, dem vom Gericht ausgesprochenen Strafmass und der gerichtlichen Anordnung einer ambulanten oder stationären Massnahme von Interesse. Aus den in den BVD gelagerten Akten der einzelnen Fälle wurden Strafregisterauszüge älteren Datums herangezogen sowie, falls vorhanden, psychiatrische Gutachten gelesen.

5.1.2 Stichprobenziehung

Wie erwähnt stellt die Erhebung einer Basisrate der Rückfälligkeit im Rahmen der retrospektiven Studie einen zentralen Bestandteil dar. Damit eine genügend lange Time-at-risk gegeben ist, musste auf ältere Fälle zurückgegriffen werden. Konkret bilden Personen die Stichprobe, für welche im Jahre 2004 oder 2005 ein sog. „Geschäftsabschluss“ bei den BVD ausgelöst wurde. Ein Geschäftsabschluss bedeutet im Regelfall, dass die Time-at-risk startet – sei dies durch eine Entlassung nach regulärem Haftende, eine bedingte Entlassung oder durch das Urteil selbst (wenn beispielsweise eine Verurteilung zu einer Geldstrafe vorlag oder eine bedingte Freiheitsstrafe ausgesprochen wurde). In der vorliegenden Studie liegt somit im Regelfall eine Time-at-risk von etwa sechs Jahren vor.

Aufgrund des besonderen Interesses an der Erkennung von potentiellen G/S-Straftäter durch das FaST als C-Fälle und deren geringem Anteil in der Gesamtheit aller Straftäter, erschien es wichtig, diese überproportional zur Verteilung in der Grundpopulation in die Stichprobe aufzunehmen. Dazu erstellte die kantonale Informatik eigens für die ROS-Ergebnisevaluation zwei Listen (sog. Crystal Reports) aus dem RIS. Diese beinhalten alle Geschäfte, die 2004 oder 2005 abgeschlossen wurden, und zeigen zudem das Urteil, welches zur Geschäftseröffnung geführt hat (sog. Anlassdelikt). Liste 1 beinhaltet alle Personen, deren Anlassdelikt durch das FaST besonders stark gewichtet wird und die somit eine erhöhte Wahrscheinlichkeit für eine C-Wertung haben. Es sind dies die im FaST-Manual aufgelisteten G/S- sowie schwerwiegenden SVG-Delikte (Bewährungs- und Vollzugsdienste Zürich, Fachsupport und Lernprogramme, 2013; siehe Auszug daraus in Anhang A, Seite 82). Liste 2 beinhaltet alle Personen, die ein anderes Anlassdelikt (keine G/S-Delikte oder schwerwiegende SVG-Delikte) begangen haben. Die erhaltenen Daten wurden in einem ersten Schritt bereinigt. Diese Bereinigung wurde aufgrund folgender Probleme notwendig:

1. Wurde für eine Person im Zeitraum von 2004 bis 2005 sowohl ein Geschäft aufgrund eines G/S- bzw. schwerwiegenden SVG-Delikttes als auch ein Geschäft aufgrund eines übrigen Delikttes abgeschlossen, erschien sie auf beiden Listen.
2. Wurden für eine Person im Zeitraum von 2004 bis 2005 entweder zwei Geschäfte aufgrund von G/S- bzw. schwerwiegenden SVG-Deliktten oder zwei Geschäfte aufgrund von übrigen Deliktten abgeschlossen, erschien sie auf einer Liste doppelt.

Im ersten Fall wurde die Person per Zufallsprinzip einer der beiden Listen zugeordnet. Im zweiten Fall wurde eines der auf einer Liste aufgeführten Geschäfte zufällig ausgewählt. Nach der Datenbereinigung ergab sich die in Tabelle 4 aufgeführte Verteilung. Da ein besonderes Interesse an C-Fällen besteht, wurde eine Vollerhebung der Liste 1 durchgeführt (540 Fälle). Aus Liste 2 wurden zufällig 193 Personen gezogen. Die Gesamtstichprobe besteht somit aus 733 Personen. Die Verschiebung des Verhältnisses der Anzahl Fälle in Liste 1 und 2 von der Grundpopulation zur Stichprobe macht es nötig, für die meisten Auswertungen einen Gewichtungsfaktor einzusetzen, der den relativen Anteil der Fälle aus Liste 1 wieder dem ursprünglichen Anteil anpasst und diese im vorliegenden Fall mit dem Faktor 0,21 („herunter“-)gewichtet (s. Tabelle 4). Da die Fälle aus Liste 2 unterrepräsentiert erhoben wurden, werden diese mit einem Faktor von 3,21 („herauf“-)gewichtet. Wird der Gewichtungsfaktor für die Berechnungen eingesetzt, wird dies an gegebener Stelle erwähnt.

Tabelle 4: Grundpopulation und Stichprobe der retrospektiven Studie

	Liste 1 G/S- bzw. SVG-Delikte ge- mäss FaST-Manual	Liste 2 Übrige Delikte	Total
Grundpopulation (von den BVD abgeschlosse- ne Fälle 2004/2005)	540 15.50%	2944 84.50%	3484 100.00%
Stichprobe	540 73.67%	193 26.33%	733 100.00%
Gewichtungsfaktor	15.50/73.67= 0.21	84.50/26.33= 3.21	

Gerade weil es sich bei den Daten der retrospektiven Studie um bereits abgeschlossene Fälle handelt, müssen gewisse Personen für einzelne Auswertungen ausgeschlossen werden. Es wurden folgende Ausschlussgründe definiert:

Tabelle 5: Ausschlussgründe

Ausschlussgrund	Anzahl ausgeschlossener Personen
1. Person gestorben	25
2. Person wurde des Landes verwiesen	24
3. Person ist untergetaucht	20
4. Person bis heute in stationärer Massnahme oder Gefängnis	3
5. Urteil zum Anlassdelikt liegt über 20 Jahre zurück	2
Gesamt	74

Dabei ist zu beachten, dass diese Personen dann ausgeschlossen werden müssen, wenn Berechnungen zur Rückfälligkeit angestellt werden, da bei ihnen keine Aussage über das Vorhandensein und die Dauer der Time-at-risk gemacht werden kann.

5.1.3 Erhebungsinstrumente

Für die Datenerhebung kamen zwei Erhebungsinstrumente (FaST und VRAG) zum Einsatz, welche nachfolgend beschrieben werden.

5.1.3.1 FaST

Für die Erhebung wurde die Version 2.01 des CR!MS verwendet. Zusätzlich zu den Ausprägungen A, B oder C werden weitere Angaben in der FaST-Ausgabe aufgeführt (Angaben zur Anlasstat, den Vorstrafen, zusätzlich problematischen Aspekten und Risikominderungsfaktoren). Eine anonymisierte Ausgabe des FaST findet sich zur Veranschaulichung in Anhang B (Seite 86).⁵

5.1.3.2 VRAG

Eine von den Autoren des VRAG autorisierte deutsche Übersetzung des Instruments (Rossegger, Urbaniok, Danielsson & Endrass, 2009) steht als online-basiertes Tool zur Verfügung, wovon im Rahmen der Ergebnisevaluation Gebrauch gemacht wurde.⁶ Der VRAG ist ein aktuarisches Prognoseinstrument, welches bei schweren G/S-Delikten angewendet werden kann. Konkret sind dies (Endrass, Rossegger & Urbaniok, 2007):

- Mord
- Versuchter Mord
- Vorsätzliche Tötung
- Versuchte vorsätzliche Tötung
- Totschlag
- Freiheitsberaubung und Entführung
- Einfache und schwere Körperverletzung
- Qualifizierter Raub
- Vergewaltigung
- Sexuelle Nötigung
- Sexuelle Handlungen mit Kindern
- Schändung
- Ausnützung der Notlage

Ein Spezialfall bildet die Brandstiftung, welche nicht wie die aufgeführten Delikte zur Begegnung und zum Körperkontakt zwischen einem Täter und einem Opfer führt. Quinsey et al. (2006) führen aus, dass bei Vorliegen einer Brandstiftung ein VRAG angewendet werden kann, dieser jedoch auch in diesem Fall die Rückfallwahrscheinlichkeit in Bezug auf ein Gewaltdelikt, und nicht auf eine erneute Brandstiftung ausgibt. In der vorliegenden Studie wurde der VRAG bei Vorliegen einer Brandstiftung angewendet.

Der VRAG besteht aus zwölf Items, die unterschiedlich stark (positiv oder negativ) gewichtet werden. Der sich daraus ergebende Summenwert bildet neun Risikoabstufungen, welche mit der Rückfallwahrscheinlichkeit eines erneuten Gewaltdelikt nach sieben und zehn Jahren assoziiert sind (vgl.

⁵ Um diese weiteren Angaben effizient in die SPSS-Maske übertragen zu können, errichtete eine Fachperson für den IT-Bereich der BVD einzig für unsere Zwecke ein Excel-basiertes Add-In, welches die Übertragung der FaST-Ausgabe in die SPSS-Maske erheblich vereinfachte.

⁶<http://www.fotres.ch/index.cfm?&content=9010&CFID=14629644&CFTOKEN=b049584fabb824d5-295A1700-13D4-FEFB-5E36623270EF8EFB>, besucht am 16.07.2013.

Tabelle 6). Der VRAG wurde immer dann durchgeführt, wenn die Person im Rahmen des Anlassdelikts oder einer Vorstrafe ein auf Seite 15 aufgeführtes Delikt begangen hatte (inkl. Brandstiftung). Die Psychopathy-Checkliste Revised (PCL-R; Hare, 1970) stellt eines der zwölf Items dar. Da zur Anwendung der PCL-R therapeutische Kenntnisse notwendig sind, wurde sie standardmässig durch die Childhood and Adolescent Taxon Scale (CATS ; Harris, Rice & Quinsey, 1994) ersetzt, wie dies von Quinsey et al. (2006) als Alternative vorgeschlagen wird.

Tabelle 6: VRAG-Risikoklassen und Rückfallwahrscheinlichkeiten

Summenwert	VRAG-Risikoklasse	Rückfallwahrscheinlichkeiten (in %)	
		Nach 7 Jahren	Nach 10 Jahren
≤ -22	1	0	8
-21 bis -15	2	8	10
-14 bis -8	3	12	24
-7 bis -1	4	17	31
0 bis +6	5	35	48
+7 bis +13	6	44	58
+14 bis +20	7	55	64
+21 bis +27	8	76	82
≥ +28	9	100	100

Quelle: Quinsey et al. (2006).

5.1.4 Beschreibung der erhobenen Variablen

Mittels oben aufgeführter Informationsquellen und Erhebungsinstrumente wurden die für die retrospektive Studie relevanten Variablen erhoben. Diese werden nachfolgend beschrieben.

5.1.4.1 Variablen zum FaST

Wie erwähnt werden in der Ausgabe des FaST zusätzlich zur Ausprägung A, B oder C Angaben zur Anlasstat, den Vorstrafen, zusätzlichen problematischen Aspekten und der Risikominderung berücksichtigt (siehe Anhang B, Seite 86). Besonders relevant für die Ergebnisevaluation sind die Angaben zum Anlassdelikt, welche gar mit weiteren Spezifizierungen erhoben werden (siehe Kapitel 5.1.4.3).

5.1.4.2 Variablen zum VRAG

Die Durchführbarkeit des VRAG hat sich als Herausforderung dargestellt. Einerseits kann der VRAG nur bei gewissen Delikten durchgeführt werden (siehe Kapitel 5.1.3.2), andererseits hat sich herausgestellt, dass die Datenlage der Fälle häufig unzureichend war, so dass kein VRAG durchgeführt werden konnte. Gerade Aussagen zu den familiären und schulischen Verhältnissen des Täters waren oft nicht in den Urteilen vorhanden. Psychiatrische Gutachten fehlten häufig trotz einem schweren Gewaltdelikt oder waren zu wenig detailliert, um die für einen VRAG nötigen Informationen zu ermitteln. Konnte jedoch ein VRAG angewendet werden, so wurden dazu folgende Variablen erhoben:

- VRAG-Risikoklasse (1-9)
- Durchführung aufgrund Anlassdelikt oder Vorstrafe
- Straftatbestand
- Datum des Urteils, welches aufgrund des VRAG-relevanten Straftatbestandes ausgesprochen wurde

5.1.4.3 Variablen zum Anlassdelikt

Zusätzlich zu den Angaben der FaST-Ausgabe wurden weitere Informationen zum Anlassdelikt erhoben. Diese Informationen stammen vorwiegend aus dem Strafregisterauszug, welcher den Akten beigelegt wurde. Konkret handelt es sich um folgende Angaben zum Anlassdelikt:

- Datum des Urteils
- Behörde, welche das Urteil aussprach
- Delikte, aufgrund welcher die Person verurteilt wurde: In Anlehnung an das FaST wurden jeweils bis zu vier G/S-Delikte sowie schwerwiegende SVG-Delikte pro Urteil erhoben (siehe Anhang A, Seite 82). Lag kein solches vor, wurde jeweils nur das erstgenannte Delikt erhoben.
- Begehungsdaten der Delikte
- Strafmass in Monaten
- Bedingter oder unbedingter Vollzug
- Massnahme angeordnet oder nicht und wenn ja, welche?

5.1.4.4 Variablen zur Time-at-risk

Für die Erhebung der Rückfälligkeit ist es von Interesse zu erfahren, wie lange eine Person die Gelegenheit hatte, rückfällig zu werden, bis sie es tatsächlich wurde. Weiter kann eruiert werden, welches Ereignis die Time-at-risk ausgelöst hat. Diese Informationen können dem RIS, in Einzelfällen unter Zuhilfenahme der Akten, entnommen werden. Als Auslösung der Time-at-risk wurden folgende Ereignisse erfasst (im RIS als sog. „Vorfälle“ deklariert):

- Bedingte Entlassung aus dem Strafvollzug
- Entlassung aus dem Strafvollzug nach regulärem Haftende
- Bedingte Entlassung aus einer stationären Massnahme
- Entlassung aus einer stationären Massnahme nach regulärem Ende
- Datum des Urteils des Anlassdelikts, wenn:
 - Umwandlung einer Freiheitsstrafe in eine Busse oder Gemeinnützige Arbeit
 - Ausspruch einer bedingten Freiheitsstrafe
 - Vollzug in Halbgefängenschaft
 - Ausspruch einer ambulanten Massnahme
 - Untertauchen einer Person
 - Weiterziehen des Urteils an nächste Instanz, sofern keine unbedingte Strafe

Das Datum des ersten Rückfalls wurde aus dem aktuellen Strafregisterauszug (alle im Sommer 2012 ausgestellt) abgeleitet. Dabei interessierte nicht das Datum des Urteils, sondern das Datum der Deliktbegehung, welche zu einem erneuten Urteil führte. Hierbei wurde darauf geachtet, dass der erste Rückfall möglicherweise erst in einem zweiten oder gar noch späteren Urteil nach dem Indexdelikt aufgeführt war. Dies geschieht, wenn ein Delikt für längere Zeit unentdeckt bleibt.

Durch die Erhebung des Starts der Time-at-risk und die Erfassung des ersten registrierten Rückfalls ist es möglich, die Time-at-risk bis zum ersten Rückfall auf den Tag genau zu berechnen. Einschränkend muss dazu gesagt werden, dass keine Aussagen über unentdeckte Rückfälle möglich sind.

5.1.4.5 Variablen zu Rückfällen

Mittels aktueller Strafregisterauszüge wurde die Rückfälligkeit der Täter untersucht. Dabei wurden bis zu acht Urteile detailliert aufgenommen. In Einzelfällen wurden mehr als acht Verurteilungen nach dem Anlassdelikt ausgesprochen, was in einer Notiz vermerkt wurde. Es wurden jeweils folgende Variablen erhoben:

- Datum des Urteils des Rückfalls
- Delikte, aufgrund welcher die Person verurteilt wurde: In Anlehnung an das FaST wurden jeweils bis zu vier G/S-Delikte sowie schwerwiegende SVG-Delikte erhoben (siehe Anhang B, Seite 82). Lag kein solches vor, wurde jeweils nur das erstgenannte Delikt erhoben.
- Begehungsdaten der Delikte
- Strafmass in Monaten
- Massnahme angeordnet oder nicht und wenn ja, welche?

Für die Auswertungen mussten die Straftatbestände kategorisiert werden. Dies erfolgte in einem ersten Schritt in Anlehnung an die Praxis des Psychiatrisch-Psychologischen Diensts des Amts für Justizvollzug Zürich (PPD), welcher Straftatbestände in eigenen Rückfallstudien folgendermassen kategorisierte.

- **Schwere Gewaltdelikte** (Tötung, Mord, Totschlag, Verleitung/Beihilfe zum Selbstmord, fahrlässige Tötung, Körperverletzung, Gefährdung des Lebens, Raufhandel, Angriff, Raub, Nötigung, Freiheitsberaubung und/oder Entführung, Geiselnahme, Gewalt und Drohung gegen Beamte)
- **Schwere Sexualdelikte** (sexuelle Handlungen mit Kindern, sexuelle Handlungen mit Abhängigen, sexuelle Nötigung, Vergewaltigung, Schändung, sexuelle Handlungen mit Anstaltspfleglingen, Ausnützen einer Notlage, Menschenhandel, sexuelle Belästigung, Inzest)
- **Sonstige G/S-Delikte** (Erpressung, Drohung, Exhibitionismus, Pornographie, Brandstiftung, gemeingefährliche Verbrechen, Förderung der Prostitution, unzulässige Ausübung der Prostitution)
- **Allgemeine Delikte** (mit den Subkategorien: Vermögensdelikte, Hausfriedensbruch, Fahrlässige Verursachung einer Feuersbrunst, Verbrechen gegen die öffentliche Gesundheit, Verstösse und Vergehen gegen das Betäubungsmittelgesetz, Verstösse und Vergehen gegen das Strassenverkehrsgesetz, Verstösse und Vergehen gegen das Aufenthalt und Niederlassung von Ausländern Gesetz bzw. Ausländergesetz, Verstösse und Vergehen gegen das Waffengesetz, Sonstiges)

Im Zwischenbericht II (Baur, Manzoni & Schwarzenegger, 2013) wurden alle Berechnungen zur Verteilung der Anlassdelikte sowie zur Rückfälligkeit vor dem Hintergrund dieser Kategorisierung gerechnet. Dies hat teilweise zu Resultaten geführt, die man nicht erwartet hätte. Bei genauerer Betrachtung wurde deutlich, dass die PPD-Kategorisierung inhaltlich nicht dem Aufbau des FaST entspricht (bspw. ist es nicht im Sinne des FaST, Gewalt und Drohung gegen Beamte als schweres Gewaltdelikt oder die unzulässige Ausübung der Prostitution als Sexualdelikt zu klassifizieren).

Es ist deshalb sinnvoll, die Auswertungen des zweiten Zwischenberichts anhand einer inhaltlich stringenteren Deliktkategorisierung vorzunehmen. Das FaST bietet eine Deliktkategorisierung von G/S-Delikten, welche sich auf den gesetzlichen Strafrahmen stützt. Es gibt jeweils drei Abstufungen der Schweregrade (G1-G3 bzw. S1-S3), welche ebenfalls in Anhang A (Seite 82) aufgeführt sind. Sämtliche Auswertungen werden anhand dieser Deliktkategorisierung vorgenommen.

5.1.4.6 Steuerbares Einkommen als Indikator für soziale Integration

Ist ein ehemaliger Straftäter sozial integriert, ist nicht nur davon auszugehen, dass er weniger rückfällig wird, sondern auch, dass er einer Arbeit nachgeht und somit ein Einkommen erwirtschaftet. Es wäre naheliegend, dass A-Fälle am wenigsten Probleme haben, sich nach Verbüßung einer Sanktion sozial zu integrieren, wohingegen dies C-Fällen am schwersten fallen dürfte. Dies dürfte aus einer Kombination von personen- und situationsbezogene Faktoren resultieren. Konkret wäre zu vermuten, dass C-Fälle (Personen, die im Rahmen einer Vorstrafe oder des Anlassdelikts Gewalt- oder Sexualdelikte begangen haben) eine auffälligere Persönlichkeitsstruktur aufweisen als A- und B-Fälle (personenbezogener Faktor, siehe z.B. Quinsey et al., 2006). Weiter verbüßen C-Fälle meist längere Freiheitsstrafen als A- und B-Fälle und dürften deshalb längere Zeit benötigen, bis sie sich wieder im Alltag zurecht finden als A- und B-Fälle (situationsbezogener Faktor).

Für die Stichprobe der retrospektiven Studie wurde das Steueramt des Kantons Zürich beauftragt, das steuerbare Einkommen (nur Kanton Zürich) für die Jahre 2005-2011 zu liefern. Es wird untersucht, wie sich das Einkommen von Personen über einen Zeitraum von sechs Jahren Time-at-risk entwickelte und inwiefern sich die Entwicklungen von A-, B- und C-Fällen voneinander unterscheiden.

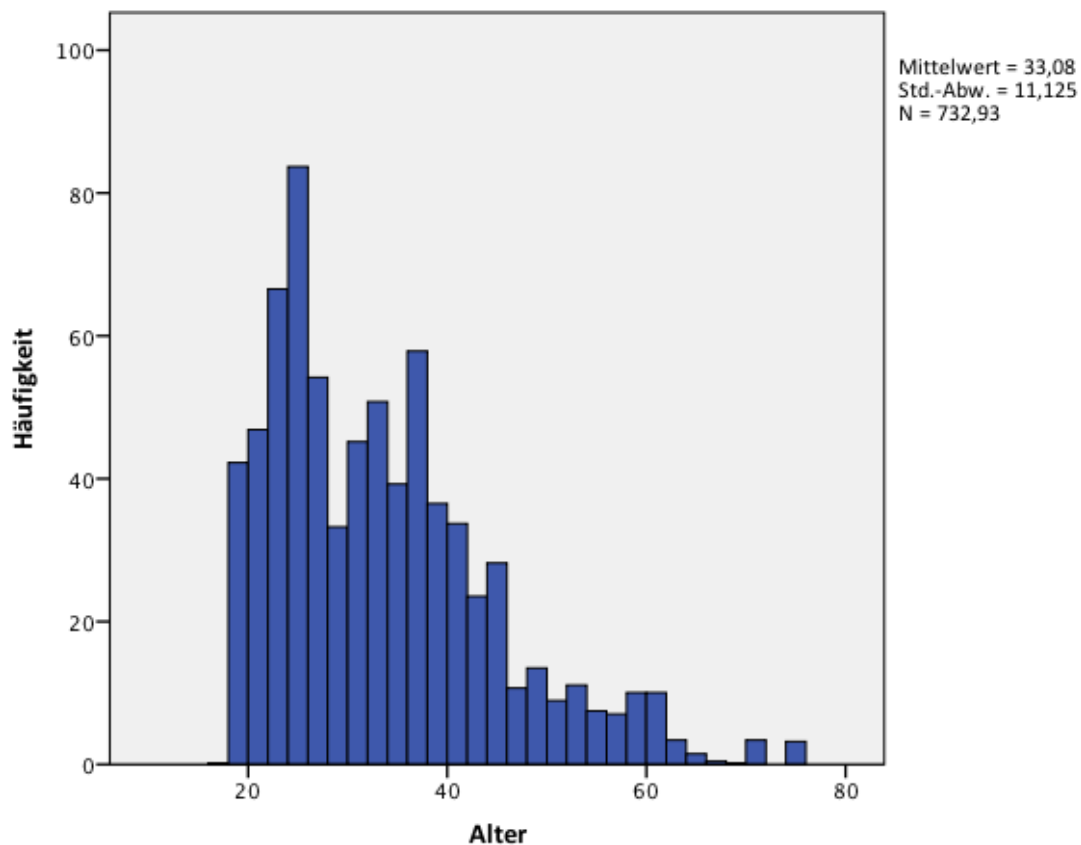
5.2 Ergebnisse

5.2.1 Beschreibung der Stichprobe

Die Verurteilten-Stichprobe besteht aus 91% Männern und 9% Frauen. Ferner sind rund 40% der Personen schweizerischer und 60% ausländischer Nationalität.

Die Personen der Stichprobe sind durchschnittlich 33.1 Jahre alt, mit einer mittleren Streuung von 11.1 Jahren. Das Alter ist hier bezogen auf die nachträgliche FaST-Anwendung im Zeitraum 2004-2005. Die Altersspanne der Verurteilten reicht von 17 bis 74 Jahren.

Abbildung 3: Stichprobe nach Alter



5.2.2 Art der Anlassdelikte nach FaST-Kategorien

Die „Art des Anlassdelikts“, das heisst, ob es sich um allgemeine Delinquenz oder um G/S-Delinquenz handelt, spielt eine zentrale Rolle für die Wertung des FaST. Es ist deshalb von Interesse, wie sich dies auf die Verteilung der FaST-Kategorien A, B und C auswirkt. Da A- und B-Fälle in der Praxis gleich behandelt werden, werden diese in den nachfolgenden Tabellen nur mit einer gestrichelten Linie voneinander getrennt, jedoch separaten Analysen unterzogen, um potentiell wichtige Unterschiede entdecken zu können. Die Anlassdelikte werden nach der neuen Deliktkategorisierung gemäss FaST-Gewichtung eingeteilt. Bei den folgenden Auswertungen wurde jeweils der Gewichtungsfaktor eingesetzt, um eine Annäherung an die Verteilung in der Grundpopulation zu erreichen. Dies führt zu einer Veränderung der Fallzahlen. Es wurden keine Ausschlüsse vorgenommen, da das Anlassdelikt für alle Personen bekannt ist.

Tabelle 7: Art des Anlassdelikts, differenziert nach FaST-Kategorien

	A	B	C	Gesamt
G1	16 36.4%	21 47.7%	7 15.9%	44 100.0%
G2	2 25.0%	1 12.5%	5 62.5%	8 100.0%
G3	0 0.0%	0 0.0%	3 100.0%	3 100.0%
S1	0 0.0%	0 0.0%	5 100.0%	5 100.0%
S2	0 0.0%	0 0.0%	0 0.0%	0 0.0%
S3	0 0.0%	0 0.0%	3 100.0%	3 100.0%
Andere Delikte	537 80.4%	94 14.1%	37 5.5%	668 100.0%
<i>Vermögensdelikt</i>	55 61.1%	32 35.6%	3 3.3%	90 100.0%
<i>Hausfriedensbruch</i>	35 79.5%	3 6.9%	6 13.6%	44 100.0%
<i>BetmG</i>	109 80.7%	16 11.9%	10 7.4%	135 100.0%
<i>Schwere SVG</i>	118 76.1%	27 17.4%	10 6.5%	155 100.0%
<i>Sonstige SVG</i>	58 93.5%	4 6.5%	0 0.0%	62 100.0%
<i>ANAG/AuG</i>	96 97.0%	0 0.0%	3 3.0%	99 100.0%
<i>Andere</i>	66 79.5%	12 14.5%	5 6.0%	83 100.0%
Gesamt	555 75.9%	116 15.9%	60 8.2%	731 100.0%

Hinweis: Unterteilung der Nicht-G/S-Delikte wird nur hier vorgenommen. Später im Bericht keine Spezifizierung der „anderen Delikte“ mehr.

Aus Tabelle 7 wird ersichtlich, dass mit Abstand die meisten Anlassdelikte als „andere Delikte“ klassifiziert werden (668 von 731 Fällen). Dieser Anteil widerspiegelt die Verteilung in der Grundpopulation, da die Daten gewichtet wurden. Weiter fällt auf, dass alle Unterkategorien der „anderen Delikte“ weitaus am häufigsten der A-Kategorie zugeteilt werden (zwischen 61.1% und 97.0%, durchschnittlich 80.4%).

Alle acht Personen, welche ein Sexualdelikt begangen haben, wurden vom FaST als C-Fall erkannt. Weniger eindeutig sieht es bei den Gewaltdelikten aus: Die meisten Personen, die ein G1-Anlassdelikt aufweisen, sind A- oder B-Fälle (36.4% bzw. 47.7%). Zwei der Personen mit einem G2-Anlassdelikt werden ebenfalls als A-Fall klassifiziert, einer als B-Fall und immerhin fünf als C-Fall. Alle drei Personen mit einem G3-Anlassdelikt sind als C-Fälle klassifiziert worden. Damit wird die Frage aufgeworfen, um welche Delikte es sich bei den G1-G3-Fällen handelt, worüber die drei nachfolgenden Tabellen Auskunft geben. Aufgrund der sehr kleinen Fallzahlen werden keine Prozentzahlen angegeben.

Tabelle 8 stellt die G1-Anlassdelikte dar. Es zeigt sich, dass einfache Körperverletzungen und Drohungen am häufigsten vertreten sind (16 bzw. 14 von 44 Fällen). Einfache Körperverletzungen resultieren am häufigsten in einem A-Fall, während Drohungen meist in einen B-Fall münden. Weitere vorkommende G1-Anlassdelikte sind Raufhandel sowie die eher schwerwiegenden Delikte nicht qualifizierter Raub und Nötigung.

Neben dem Anlassdelikt sind ebenfalls andere Faktoren für die FaST-Ausprägung relevant, welche der anonymisierten FaST-Ausgabe in Anhang B (Seite 86) entnommen werden können. Dies erklärt, weshalb dasselbe Anlassdelikt in unterschiedlichen FaST-Kategorien resultieren kann. Sind jedoch keine oder nur wenige andere gewichtige Faktoren vorhanden, so münden G1-Delikte wohl vornehmlich in der A-Kategorie.

Tabelle 8: G1-Anlassdelikte, differenziert nach FaST-Kategorien

	A	B	C	Gesamt
Einfache Körperverletzung (Art. 123 StGB)	9	4	3	16
Raufhandel (Art. 133 StGB)	1			1
Raub, nicht qualifiziert (Art. 140 Abs. 1 StGB)	1	6	2	9
Drohung (Art. 180 StGB)	4	8	1	14
Nötigung (Art. 181 StGB)	1	3	1	5
Gesamt	16	21	7	44

Tabelle 9 zeigt, dass (nach Gewichtung) insgesamt acht G2-Delikte vorhanden waren. Es wird ersichtlich, dass selbst G2-Anlassdelikte in der A-Kategorie resultieren können. Dies wird dann der Fall sein, wenn keine oder nur wenige andere für das FaST gewichtige Faktoren vorhanden sind (vgl. Anhang B, Seite 86). Unter dem Vorbehalt der sehr kleinen Fallzahlen ist jedoch ersichtlich, dass G2-Fälle am häufigsten zu einer C-Kategorisierung führen.

Tabelle 9: G2-Anlassdelikte, differenziert nach FaST-Kategorien

	A	B	C	Gesamt
Gefährdung des Lebens (Art. 129 StGB)	1		1	2
Angriff (Art. 134 StGB)	1			1
Raub, qualifiziert (Art. 140 Abs. 2 oder 3 StGB)		1	2	3
Freiheitsberaubung/Entführung (Art. 183 StGB)			1	1
Brandstiftung (Art. 221 StGB)			1	1
Gesamt	2	1	5	8

Tabelle 10 führt schliesslich die G3-Delikte auf, welche sodann auch alle als C-Fall kategorisiert werden. Nach der Gewichtung waren lediglich drei Personen mit einem G3-Anlassdelikt vorhanden (eine Tötung, eine schwere Körperverletzung, ein qualifizierter Raub).

Tabelle 10: G3-Anlassdelikte, differenziert nach FaST-Kategorien

	A	B	C	Gesamt
Tötung (Art. 111 StGB)			1	1
Schwere Körperverletzung (Art. 122 StGB)			1	1
Raub, qualifiziert (Art. 140 Abs. 4 StGB)			1	1
Gesamt			3	3

Insgesamt kann gesagt werden, dass der Anteil an C-Fällen von G1-Anlassdelikten bis G3-Anlassdelikten stetig zunimmt, womit gezeigt wird, dass dem Anlassdelikt bei der Kategorisierung des FaST eine zentrale Bedeutung zukommt.

5.2.3 Art der Rückfälligkeit nach FaST-Kategorien

Zur Bestimmung der Güte des FaST interessiert insbesondere, in welchem Zusammenhang die FaST-Kategorien mit der „Art der Rückfälligkeit“ (allgemeine Delinquenz oder G/S-Delinquenz) stehen. Dabei wären separate Auswertungen für die Abstufungen der „Schwere“ von Gewalt- und Sexualdelikten (G1-G3 bzw. S1-S3) zwar wünschenswert, die Durchführung statistischer Tests ist jedoch aufgrund kleiner Fallzahlen nicht möglich.⁷ Gewalt und Sexualdelikte werden deshalb nachfolgend zusammengefasst und durch den Begriff G/S-Delinquenz ausgedrückt. Bezüglich des Zusammenhangs der Art der Rückfälligkeit und der FaST-Kategorisierung wird eine Annahme getroffen, welche sich an der FaST-Konzeption ausrichtet und anhand Tabelle 11 erläutert wird.

Tabelle 11: Kombinationsmöglichkeiten von Rückfallarten und FaST-Kategorien

	A	B	C
Kein Rückfall oder Rückfall allgemeine Delinquenz	Treffer	Treffer	Falsch positiv
Rückfall G/S-Delinquenz	Falsch negativ	Falsch negativ	Treffer

Mit der Ausgabe C zeigt das FaST einen erhöhten Abklärungsbedarf bezüglich G/S-Delinquenz auf. Implizit daraus abgeleitet bedeutet dies, dass C-Fälle ein erhöhtes Risiko aufweisen, mit einem G/S-Delikt rückfällig zu werden. Es wird somit erwartet, dass C-Fälle gehäuft mit einem G/S-Delikt rückfällig werden (Treffer). A- und B-Fälle hingegen weisen keinen erhöhten Abklärungsbedarf bzw. einen erhöhten Abklärungsbedarf bezüglich allgemeiner Delinquenz auf, weshalb sie seltener mit einem G/S-Delikt rückfällig werden sollten (falsch negativ).

Annahme:

C-Fälle werden häufiger mit einem G/S-Delikt rückfällig als A- und B-Fälle (Treffer).

Um die Annahme zu prüfen, wurde ein χ^2 -Test durchgeführt. Ergibt ein χ^2 -Test ein signifikantes Resultat, kann mittels der sog. standardisierten Residuen eruiert werden, welche Kombination der Ausprägung zweier Variablen (hier: FaST und Rückfallart) signifikant häufiger oder seltener als andere Ausprägungskombinationen vorkommt.⁸

Für die Analysen wurde der Gewichtungsfaktor eingesetzt. Personen, welche einen Ausschlussgrund gemäss Tabelle 5 erfüllen, flossen nicht in die Auswertungen ein, da für sie keine Time-at-risk ermittelt werden konnte.

⁷ Deskriptive Darstellungen der Schwere von Gewaltrückfällen folgen in Kapitel 5.2.4.

⁸ Im Rahmen eines χ^2 -Tests wird bei Kreuztabellen (z.B. Tabelle 11) eine Schätzung der Anzahl an Fällen pro Zelle vorgenommen. Diese Schätzung basiert auf der Spaltensumme, der Zeilensumme und der Anzahl berücksichtigter Fälle. Die beobachtete Anzahl an Fällen wird dann mit der erwarteten Anzahl an Fällen verglichen. Daraus ergibt sich ein sog. Residualwert, welcher anzeigt, ob diese signifikant voneinander abweichen (signifikante Abweichung wenn Residualwert < -1.96 bzw. > 1.96) (Field, 2013).

Tabelle 12 zeigt die Verteilung der Rückfälle differenziert nach FaST-Kategorien auf. Es besteht ein signifikanter Zusammenhang zwischen der FaST-Kategorie und der Rückfallart.⁹ Es zeigt sich, dass C-Fälle signifikant häufiger mit einem G/S-Delikt rückfällig wurden als A- und B-Fälle (Residualwert von 2.7, 25% aller C-Fälle), womit die Annahme bestätigt werden konnte.

Tabelle 12: Rückfallart, differenziert nach FaST-Kategorien

		A	B	C	Gesamt
Kein Rückfall oder Rückfall allgemeine Delinquenz	<i>Absolute Häufigkeiten:</i>	430	86	39	555
	<i>Spaltenprozenze:</i>	90.0%	86.0%	75.0%	88.1%
	<i>Residualwerte:</i>	0.4	-0.2	-1.0	
Rückfall G/S-Delinquenz	<i>Absolute Häufigkeiten:</i>	48	14	13	75
	<i>Spaltenprozenze:</i>	10.0%	14.0%	25.0%	11.9%
	<i>Residualwerte:</i>	-1.2	0.6	2.7	
Gesamt	<i>Absolute Häufigkeiten:</i>	478	100	52	630
	<i>Spaltenprozenze:</i>	100.0%	100.0%	100.0%	100.0%

Trotzdem werden auch 10.0% der A-Fälle und 14.0% der B-Fälle mit einem G/S-Delikt rückfällig (Tabelle 12). Gemäss FaST-Konzeption könnte es sich hierbei um Personen handeln, welche zum Zeitpunkt der FaST-Einschätzung kein G/S-Delikt aufwiesen (weder im Rahmen des Anlassdelikts, noch im Rahmen von Vorstrafen), denn die Art des Anlassdelikts bzw. der Vorstrafen stellen wichtige Gewichtungsfaktoren für das FaST dar. Um dies zu überprüfen werden nur jene 75 Personen der Tabelle 12 berücksichtigt, welche mit einem G/S-Delikt rückfällig wurden.

Tabelle 13 zeigt, dass die beiden Variablen „FaST-Kategorie“ und „G/S-Delikt im Moment der FaST-Einschätzung“ einen hochsignifikanten Zusammenhang aufweisen.¹⁰ Allgemein kann gesagt werden, dass es sich bei der Mehrheit der Personen, die mit einem G/S-Delikt rückfällig wurden, um Personen handelt, die zum Zeitpunkt der FaST-Einschätzung kein G/S-Delikte aufwiesen (45 Fälle, 60.0%).

Die standardisierten Residuen zeigen, dass C-Fälle, bei denen ein G/S-Delikt im Moment der FaST-Einschätzung bekannt war, häufiger mit einem G/S-Delikt rückfällig werden (nämlich alle 13 Fälle, Residualwert 3.4), während dies bei A-Fällen seltener vorkommt (7 von 48 Fällen, Residualwert -2.8). Die 10 B-Fälle stellen hingegen (knapp) keine signifikante Häufung dar.

Es kann gefolgert werden, dass ein G/S-Delikt im Rahmen einer Vorstrafe oder eines Anlassdelikts durch das FaST so stark gewichtet wird, dass es meistens zu einer C-Kategorisierung führt. Das Resultat wirft jedoch auch die Frage auf, wie das Risiko bezüglich eines künftigen G/S-Delikts erkannt werden kann, wenn der Täter in Bezug auf diese Deliktkategorie zum Zeitpunkt der FaST-Beurteilung noch nicht mit einem G/S-Delikt strafrechtlich registriert wurde (siehe dazu auch die Ausführungen in Kapitel 8.1). Dies ist umso bedeutender, als dass dies mehr als die Hälfte der Personen, die mit einem G/S-Delikt rückfällig wurden (45 von 75 Personen), sind und diese überwiegend als A-Fall klassifiziert wurden.

⁹ $\chi^2(2) = 10.5, p \leq .01$

¹⁰ $\chi^2(2) = 38.18, p \leq .001$

Tabelle 13: Vorliegen eines G/S-Delikts (Vorstrafe oder Anlassdelikt), differenziert nach FaST-Kategorien

		A	B	C	Gesamt
Kein G/S-Delikt im Moment der FaST-Einschätzung bekannt	<i>Absolute Häufigkeiten:</i>	41	4	0	45
	<i>Spaltenprozente:</i>	85.4%	28.6%	0.0%	60.0%
	<i>Residualwerte:</i>	2.3	-1.5	-2.8	
G/S-Delikt im Moment der FaST-Einschätzung bekannt	<i>Absolute Häufigkeiten:</i>	7	10	13	30
	<i>Spaltenprozente:</i>	14.6%	71.4%	100.0%	40.0%
	<i>Residualwerte:</i>	-2.8	1.9	3.4	
Gesamt	<i>Absolute Häufigkeiten:</i>	48	14	13	75
	<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%

Anmerkung: Einbezogen wurden jene 75 Personen, die mit einem G/S-Delikt rückfällig wurden.

5.2.4 Schwere von Gewaltrückfällen nach FaST-Kategorien

Wie erwähnt lassen die kleinen Fallzahlen von G/S-Rückfällen keine statistischen Tests über die Zusammenhänge der Schwere von G/S-Delikten (G1-G3, S1-S3) mit dem FaST zu. An dieser Stelle soll jedoch trotzdem rein deskriptiv gezeigt werden, wie sich die 72 G-Rückfälle (die drei S-Rückfälle werden ausgeklammert) auf die FaST-Kategorien verteilen. Aus Tabelle 14 wird ersichtlich, dass mit Abstand am meisten G-Rückfälle der G1-Kategorie zugeordnet werden (84.8% aller Fälle). Von diesen sind wiederum 73.8% in der A-Kategorie zu finden.

Tabelle 14: Schwere von Gewaltrückfällen, differenziert nach FaST-Kategorien

		A	B	C	Gesamt
G1	<i>Absolute Häufigkeiten:</i>	45	10	6	61
	<i>Spaltenprozente:</i>	95.8%	71.4%	54.5%	84.8%
	<i>Zeilenprozente:</i>	73.8%	16.4%	9.8%	100.0%
G2	<i>Absolute Häufigkeiten:</i>	1	4	1	6
	<i>Spaltenprozente:</i>	2.1%	28.6%	9.1%	8.3%
	<i>Zeilenprozente:</i>	16.7%	66.7%	16.6%	100.0%
G3	<i>Absolute Häufigkeiten:</i>	1	0	4	5
	<i>Spaltenprozente:</i>	2.1%	0.0%	36.4%	6.9%
	<i>Zeilenprozente:</i>	20.0%	0.0%	80.0%	100.0%
Gesamt	<i>Absolute Häufigkeiten:</i>	47	14	11	72
	<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%
	<i>Zeilenprozente:</i>	65.3%	19.4%	15.3%	100.0%

A-Fälle werden überwiegend mit einem G1-Delikt rückfällig, jedoch nur sehr selten mit einem G2- oder G3-Delikt. Das FaST erkennt somit für Personen, die mit einem G1-Delikt rückfällig werden, selten einen erhöhten Abklärungsbedarf bezüglich G/S-Delinquenz (nur 9.8% sind C-Fälle). Jedoch scheint es den erhöhten Abklärungsbedarf jener Personen zu erkennen, die später mit einem G3-Delikt rückfällig werden (vier von fünf sind C-Fälle). Dieses Resultat scheint in der Tendenz mit empirischen Befunden einher zu gehen, welche für schwere Gewalttäter ein erhöhtes einschlägiges Rückfallrisiko aufzeigen (Andrews & Bonta, 2010).

5.2.5 Mittlere Time-at-risk bis zum Rückfall nach FaST-Kategorien

Von Interesse ist nicht nur, ob ein Rückfall geschieht, sondern auch, wie lange es bis zum Rückfall dauert. In Bezug auf die Evaluation des FaST ist es weiter von Interesse, ob sich die mittleren Time-at-risks bis zum Rückfall je nach FaST-Kategorie voneinander unterscheiden. Tabelle 15 zeigt die mittleren Time-at-risks in Monaten nach FaST-Kategorie auf.¹¹

Die Werte werden separat für alle Rückfälle sowie G/S-Rückfälle aufgezeigt. Eine separate Analyse für Rückfälle allgemeiner Delinquenz ist leider nicht möglich, da ein allgemeiner Rückfall nur dann erhoben wurde, wenn die Person im gleichen Urteil nicht ebenfalls wegen eines G/S-Delikts verurteilt wurde. Zudem ist es aufgrund der kleinen Fallzahlen nicht möglich, die G/S-Delikte nach ihrem Schweregrad weiter zu differenzieren. Die Fälle wurden gewichtet und die Ausschlüsse gemäss Tabelle 5 vorgenommen.

Die Medianwerte zeigen, dass C-Fälle mit acht Monaten Time-at-risk am schnellsten rückfällig werden, gefolgt von A-Fällen mit elf Monaten und B-Fällen mit sechzehn Monaten. Post-hoc Tests (Games-Howell-Prozedur), welche in der Folge einer signifikanten Varianzanalyse¹² gerechnet wurden, zeigen, dass sich die Mittelwerte der B- und C-Kategorie signifikant voneinander unterscheiden (nicht jedoch jene zwischen A- und B-Fällen bzw. A- und C-Fällen).

Betrachtet man die mittleren Time-at-risks bis zum ersten G/S-Rückfall fällt auf, dass es insbesondere bei C-Fällen länger dauert, bis ein solches verübt wird als irgendein Delikt (Mediane von 16.5 und 8 Monaten). Vergleicht man die Time-at-risk bis zum ersten G/S-Rückfall nach FaST-Kategorie fällt auf, dass diese einander ähnlich sind (Mediane von 17, 16 und 16.5 Monaten). Dementsprechend gibt die Varianzanalyse auch keine signifikanten Unterschiede aus.¹³

Zunächst einmal ist dieses Resultat kontraintuitiv, da man davon ausgeht, dass C-Fälle das höchste Rückfallrisiko bezüglich G/S-Delinquenz aufweisen und somit auch schneller mit einem solchen Delikt rückfällig werden sollten als A- und B-Fälle. Eine mögliche Erklärung für das Resultat wäre, dass C-Fälle nach Verbüßung der Sanktion stärker begleitet werden als A- und B-Fälle und so schnelle Rückfälle vermieden werden können (z.B. ambulante Therapien, dichte Betreuung durch die Bewährungshilfe, Lernprogramme). Um dies zu klären, ist aber weitere Forschung nötig.

Tabelle 15: Mittlere Time-at-risks in Monaten bis zur Rückfälligkeit, differenziert nach FaST-Kategorien

		A	B	C	Gesamt
Alle Rückfälle	<i>n:</i>	299	65	38	402
	<i>Median:</i>	11.0	16.0	8.0	11.0
	<i>Mittelwert:</i>	19.1	25.1	13.4	19.5
Rückfall G/S-Delinquenz	<i>n:</i>	48	14	13	75
	<i>Median:</i>	17.0	16.0	16.5	16.0
	<i>Mittelwert:</i>	21.3	23.2	26.5	22.5

¹¹ Als „mittlere“ Time-at-risk kommen verschieden Masse der zentralen Tendenz in Frage. Der Mittelwert hat den Nachteil, dass er anfällig auf Verzerrungen durch Extremwerte ist. Dieses Problem besteht weniger beim Medianwert. Um ihn zu berechnen, wird der obere Teil aller Werte vom unteren Teil so getrennt, dass beidseitig des Medians gleich viele Werte stehen. Statistische Tests basieren meist auf dem Mittelwert. Die Abweichungen von Mittelwert und Median sollten jedoch zumindest betrachtet werden, um die Existenz von Extremwerten abschätzen zu können (Gerrig & Zimbardo, 2008)

¹² $F(2, 398) = 3.20, p \leq .05$

¹³ $F(2, 72) = .32, p = n.s.$

5.2.6 Basisraten der Rückfälligkeit nach FaST-Kategorien und Time-at-risks

Für einen Vergleich der Rückfälligkeit von prä-ROS-Fällen und ROS-Fällen wurde im Rahmen der retrospektiven Studie eine Basisrate der Rückfälligkeit erhoben. Dabei wurden Time-at-risks von einem, drei und sechs Jahren berücksichtigt.

Tabelle 16 zeigt, wie sich die Rückfälligkeit über die Jahre entwickelte, wobei nach FaST-Kategorien differenziert wurde. Für diese Übersicht wurde die Rückfallart nicht berücksichtigt. Es interessierte lediglich, wie viele Personen in der jeweiligen Time-at-risk mindestens einmal rückfällig wurden. Spezifizierungen der Rückfallart werden ab Tabelle 17 vorgenommen. Jene Personen, die einen Ausschlussgrund gemäss Tabelle 5 aufweisen oder die entsprechende Time-at-risk nicht erreichten¹⁴, fliessen nicht in die Analysen ein. Weiter wurden die Fälle gewichtet.

Tabelle 16: Rückfälligkeit nach verschiedenen Time-at-risks, differenziert nach FaST-Kategorien

		1 Jahr	3 Jahre	6 Jahre
A	Kein Rückfall	316 (66.0%)	223 (46.6%)	197 (41.4%)
	Mind. 1 Rückfall	163 (34.0%)	256 (53.4%)	279 (58.6%)
B	Kein Rückfall	73 (73.0%)	51 (51.0%)	39 (40.2%)
	Mind. 1 Rückfall	27 (27.0%)	49 (49.0%)	58 (59.8%)
C	Kein Rückfall	24 (45.3%)	17 (32.7%)	12 (25.0%)
	Mind. 1 Rückfall	29 (54.7%)	35 (67.3%)	36 (75.0%)
Gesamt	Kein Rückfall	413 (65.3%)	291 (46.1%)	248 (39.8%)
	Mind. 1 Rückfall	219 (34.7%)	340 (53.9%)	373 (60.1%)
	<i>Total</i>	<i>632 (100.0%)</i>	<i>631 (100.0%)</i>	<i>621 (100.0%)</i>

Tabelle 16 zeigt, dass nach einem Jahr 34.7% aller Personen rückfällig wurden. Bei den A-Fällen sind es 34.0%, bei den B-Fällen 27.0% und bei den C-Fällen 54.7%. Somit werden C-Fälle deutlich schneller rückfällig als A- und B-Fälle. Weiter ist ersichtlich, dass während sechs Jahren Time-at-risk mehr C-Fälle als A- und B-Fälle rückfällig werden (75.0% aller C-Fälle versus 58.6% der A-Fälle und 59.8% der B-Fälle).

Tabelle 17 bis Tabelle 19 zeigen die Basisraten der Rückfälligkeit nach den verschiedenen Time-at-risks und differenzieren nach Rückfallart („kein Rückfall“, „Rückfall allgemeine Delinquenz“ und „Rückfall G/S-Delinquenz“). Separate Analysen für Gewalt- und Sexualrückfälle sind aufgrund der kleinen Fallzahlen wiederum nicht möglich. Ob Zusammenhänge zwischen den FaST-Kategorien und der Rückfallart bestehen, wird mittels χ^2 -Tests bzw. exakter Tests nach Fisher eruiert.¹⁵

¹⁴ Eine Person erreichte die Time-at-risk von drei Jahren nicht, elf Personen erreichten die Time-at-risk von sechs Jahren nicht.

¹⁵ Da ein χ^2 -Test gewisse Voraussetzungen an die Daten stellt und eine dieser Voraussetzungen bei einer Time-at-risk von 1 Jahr nicht erfüllt war, wurde stattdessen ein exakter Test nach Fisher gerechnet, welcher bei Nichterfüllung von Voraussetzungen anstelle eines χ^2 -Tests angewendet werden kann. In der Folge eines signifikanten exakten Tests nach Fisher dürfen jedoch die standardisierten Residuen nicht interpretiert werden, weshalb es nicht möglich ist zu eruieren, welche Kombinationen von Ausprägungen zweier Variablen überzufällig selten oder häufig vorkommen.

Tabelle 17: Rückfälligkeit, differenziert nach FaST-Kategorien, 1 Jahr Time-at-risk

			A	B	C	Gesamt
1 Jahr Time-at-risk	Kein Rückfall	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	316 66.0%	73 73.0%	24 45.3%	413 65.3%
	Rückfall allgemeine Delinquenz	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	147 30.7%	22 22.0%	23 43.4%	192 30.4%
	Rückfall G/S-Delinquenz	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	16 3.3%	5 5.0%	6 11.3%	27 4.3%
	Gesamt	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	479 100.0%	100 100.0%	53 100.0%	632 100.0%

$p \leq .01$, exakter Test nach Fisher.

Tabelle 17 zeigt den signifikanten Zusammenhang von FaST-Kategorie und Rückfallart auf. Es wird ersichtlich, dass C-Fälle nach einem Jahr sowohl bezüglich der allgemeinen Delinquenz als auch bezüglich der G/S-Delinquenz häufiger rückfällig werden als A- und B-Fälle: 43.4% aller C-Fälle versus 30.7% bzw. 22.0% aller A- bzw. B-Fälle werden innerhalb eines Jahres mit einem *allgemeinen Delikt* rückfällig und 11.3% aller C-Fälle versus 3.3% bzw. 5.0% aller A- bzw. B-Fälle werden innerhalb eines Jahres mit einem *G/S-Delikt* rückfällig.

Nach drei (Tabelle 18) und nach sechs Jahren (Tabelle 19) zeigt sich das gleiche Muster wie nach einem Jahr: C-Fälle werden häufiger mit einem allgemeinen Delikt und einem G/S-Delikt rückfällig als A- und B-Fälle. Der Zusammenhang ist jedoch nach drei Jahren knapp nicht signifikant. Nach sechs Jahren hingegen ist er signifikant und die standardisierten Residuen zeigen, dass die signifikante Merkmalskombination bei den C-Fällen liegt, die gehäuft mit einem G/S-Delikt rückfällig werden.

Tabelle 18: Rückfälligkeit, differenziert nach FaST-Kategorien, 3 Jahre Time-at-risk

			A	B	C	Gesamt
3 Jahre Time-at-risk	Kein Rückfall	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	223 46.6%	51 51.0%	17 32.7%	291 46.1%
	Rückfall allgemeine Delinquenz	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	218 45.5%	36 36.0%	27 51.9%	281 44.5%
	Rückfall G/S-Delinquenz	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	38 7.9%	13 13.0%	8 15.4%	59 9.4%
	Gesamt	<i>Absolute Häufigkeiten:</i> <i>Spaltenprozente:</i>	479 100.0%	100 100.0%	52 100.0%	631 100.0%

$\chi^2(4) = 9.42$, $p = n.s.$

Tabelle 19: Rückfälligkeit, differenziert nach FaST-Kategorien, 6 Jahre Time-at-risk

			A	B	C	Gesamt
6 Jahre Time-at-risk	Kein Rückfall	<i>Absolute Häufigkeiten:</i>	197	39	12	248
		<i>Spaltenprozente:</i>	41.4%	40.2%	25.0%	39.8%
		<i>Residualwerte:</i>	0.5	0.0	-1.6	
	Rückfall allgemeine Delinquenz	<i>Absolute Häufigkeiten:</i>	231	44	25	300
		<i>Spaltenprozente:</i>	48.5%	45.4%	52.1%	48.3%
		<i>Residualwerte:</i>	0.1	-0.4	0.4	
	Rückfall G/S-Delinquenz	<i>Absolute Häufigkeiten:</i>	48	14	11	73
		<i>Spaltenprozente:</i>	10.1%	14.4%	22.9%	11.8%
		<i>Residualwerte:</i>	-1.1	0.8	2.3	
	Gesamt	<i>Absolute Häufigkeiten:</i>	476	97	48	621
		<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%

 $\chi^2(4) = 10.06, p \leq .05$

Diese Resultate deuten darauf hin, dass das FaST jene Fälle ausfindig macht (d.h. als C-Fälle klassifiziert), die gefährdet sind mit einem G/S-Delikt rückfällig zu werden. Die Konzeption des FaST, wonach jene Fälle eruiert werden sollen, die einen besonderen Abklärungsbedarf zur Verhinderung von G/S-Rückfällen aufweisen, widerspiegelt sich somit in den Befunden.

Bezüglich der B-Fälle ist jedoch anhand der vorliegenden Resultate keine Übereinstimmung mit der FaST-Konzeption festzustellen. Diese sollten nämlich auf einen erhöhten Abklärungsbedarf (und indirekt auf ein erhöhtes Rückfallrisiko) hinsichtlich allgemeiner Delinquenz hinweisen, was durch die Werte in Tabelle 17 bis Tabelle 19 nicht gestützt wird. Dies könnte jedoch auf eine fehlende Differenzierung der allgemeinen Delinquenz zurückzuführen sein. So wurden alle Delikte, die gemäss FaST-Konzeption nicht als Gewalt oder Sexualdelikt deklariert wurden (siehe Anhang A, Seite 82) als allgemeine Delikte kategorisiert. Es müsste separat geprüft werden, ob eine B-Kategorisierung durch FaST die allgemeine Delinquenz besser vorhersagt, wenn man gewisse Delikte (z.B. Verstösse gegen das AuG) aus der Definition der allgemeinen Delinquenz ausklammern würde.

5.2.7 Time-at-risk-auslösendes Ereignis nach FaST-Kategorien

Das FaST gibt mit seiner Klassifikation den Abklärungsbedarf einer straffälligen Person an. Das *Urteil* sowie allfällige *Vollzugslockerungsstufen* orientieren sich am eruierten Interventionsbedarf.¹⁶ Da der Abklärungs- und der Interventionsbedarfs eng miteinander verknüpft sind (durch eine Abklärung wird ein Interventionsbedarf bestimmt) lohnt es sich, den Zusammenhang zwischen der FaST-Kategorie und dem Ereignis, welches die Time-at-risk auslöste (dabei kann es sich um das Urteil selbst oder um Vollzugslockerungsstufen handeln), genauer zu betrachten. Nachfolgend werden die bereits in Kapitel 5.1.4.4 erläuterten Auslösungsgründe der Time-at-risk nochmals aufgelistet:

- Bedingte Entlassung aus dem Strafvollzug
- Entlassung aus dem Strafvollzug nach regulärem Haftende
- Bedingte Entlassung aus einer stationären Massnahme
- Entlassung aus einer stationären Massnahme nach regulärem Ende
- Datum des Urteils des Anlassdelikts, wenn:
 - Umwandlung einer Freiheitsstrafe in Gemeinnützige Arbeit
 - Ausspruch einer bedingten Freiheitsstrafe
 - Vollzug in Halbgefängenschaft
 - Ausspruch einer ambulanten Massnahme
 - Untertauchen einer Person
 - Umwandlung einer Freiheitsstrafe in eine Busse
 - Weiterziehen des Urteils an nächste Instanz, sofern keine unbedingte Strafe

Zur bedingten Entlassung ist zu spezifizieren, dass diese gemäss Art. 86 StGB nur dann zur Anwendung kommt, wenn das Strafmass mindestens drei Monate beträgt und es das Verhalten einer Person im Strafvollzug rechtfertigt, sie bedingt zu entlassen. Es darf nicht angenommen werden, die Person werde weitere Verbrechen oder Vergehen begehen.

Ebenfalls im Rahmen einer stationären Massnahme erfolgt eine bedingte Entlassung, sobald es der Zustand der Person rechtfertigt, dass ihr Gelegenheit gegeben wird, sich in der Freiheit zu bewähren (Art. 62 StGB).

Sind die Voraussetzungen für eine bedingte Entlassung nicht gegeben, erfolgt die Entlassung nach regulärem Haftende bzw. nach Erreichen der Höchstdauer einer stationären Massnahme nach Art. 60 und 61 StGB.

Neben der bedingten Entlassung aus dem Strafvollzug oder einer stationären Massnahme sowie der regulären Entlassung aus denselben gibt es auch die Möglichkeit, dass die Time-at-risk einer Person direkt nach Urteilsausspruch beginnt. Dies ist dann der Fall, wenn eine der sechs oben aufgeführten Optionen auftritt.

Nachfolgende Kreuztabelle zeigt auf, wie sich die Time-at-risk-auslösenden Ereignisse auf die FaST-Kategorien verteilen. Es werden keine Ausschlüsse vorgenommen, da alle Personen einer FaST-Kategorie zugeordnet werden können sowie der Grund der Auslösung der Time-at-risk bei allen bekannt ist. Die Daten werden jedoch gewichtet, um eine Annäherung der Verteilung in der Grundpopulation zu erreichen.

¹⁶ Für eine Person mit erhöhtem Interventionsbedarf wird beispielsweise eher eine Massnahme angeordnet (Urteil) bzw. es wird beispielsweise keine bedingte Entlassung aus dem Strafvollzug gewährt (Vollzugslockerungsstufe)

Die Personen, welche nach regulärem Haftende entlassen wurden, werden in zwei Kategorien eingeteilt: Eine Kategorie enthält Personen, deren Strafmass weniger als drei Monate betrug, die andere enthält jene Personen, deren Strafmass mindestens 3 Monate betrug. Dies aus dem bereits erwähnten Grund, dass Personen mit einer kurzen Freiheitsstrafe von weniger als drei Monaten nach Art. 86 StGB nicht die Möglichkeit auf eine bedingte Entlassung haben.

Tabelle 20: Ereignisse, welche die Time-at-risk auslösten, differenziert nach FaST-Kategorien

		A	B	C	Gesamt
Entlassung nach regulärem Haftende (Strafmass weniger als 3 Monate)	<i>Absolute Häufigkeiten:</i>	67	11	7	85
	<i>Spaltenprozente:</i>	12.0%	9.4%	11.9%	11.6%
	<i>Zeilenprozente:</i>	78.8%	12.9%	8.3%	100.0%
Entlassung nach regulärem Haftende (Strafmass mindestens 3 Monate)	<i>Absolute Häufigkeiten:</i>	26	18	14	58
	<i>Spaltenprozente:</i>	4.7%	15.4%	23.7%	7.9%
	<i>Zeilenprozente:</i>	44.8%	31.0%	24.2%	100.0%
Bedingte Entlassung aus dem Strafvollzug	<i>Absolute Häufigkeiten:</i>	130	29	19	178
	<i>Spaltenprozente:</i>	23.3%	24.8%	32.2%	24.3%
	<i>Zeilenprozente:</i>	73.0%	16.3%	10.7%	100.0%
Entlassung aus stationärer Massnahme nach regulärem Ende	<i>Absolute Häufigkeiten:</i>	7	10	4	21
	<i>Spaltenprozente:</i>	1.3%	8.5%	6.8%	2.9%
	<i>Zeilenprozente:</i>	33.3%	47.7%	19.0%	100.0%
Bedingte Entlassung aus stationärer Massnahme	<i>Absolute Häufigkeiten:</i>	0	7	1	8
	<i>Spaltenprozente:</i>	0.0%	6.0%	1.7%	1.1%
	<i>Zeilenprozente:</i>	0.0%	87.5%	12.5%	100.0%
Urteil des Anlassdelikts	<i>Absolute Häufigkeiten:</i>	327	42	14	383
	<i>Spaltenprozente:</i>	58.7%	35.9%	23.7%	52.2%
	<i>Zeilenprozente:</i>	85.3%	11.0%	3.7%	100.0%
Gesamt	<i>Absolute Häufigkeiten:</i>	557	117	59	733
	<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%
	<i>Zeilenprozente:</i>	76.0%	16.0%	8.0%	100.0%

Tabelle 20 zeigt, dass die Time-at-risk bei den meisten Fällen (52.2%) durch das Anlassdelikt ausgelöst wurde. Von diesen sind 85.3% A-Fälle, 11.0% B-Fälle und 3.7% C-Fälle. Das zweithäufigste Ereignis, welches die Time-at-risk auslöste, war die bedingte Entlassung aus dem Strafvollzug (24.3%). Wiederum werden A-Fälle mit 73.0% häufiger bedingt entlassen als B- und C-Fälle (16.3% bzw. 10.7%). Entlassungen nach regulärem Haftende sind seltener, sei dies im Rahmen von kurzen Freiheitsstrafen (11.6%) oder bei Vorliegen eines Strafmasses von mindestens drei Monaten (7.9%). Entlassungen aus stationären Massnahmen (bedingt oder nach regulärem Ende), sind sehr selten (1.1% bzw. 2.9%).

Für die Evaluation des FaST scheint insbesondere relevant, wie sich Personen, welche bedingt versus nach regulärem Haftende entlassen wurden, auf die FaST-Kategorien verteilen. Es kann davon ausgegangen werden, dass Personen mit eher geringem Abklärungsbedarf die Voraussetzungen für eine bedingte Entlassung erfüllen, während dies bei Personen mit erhöhtem Abklärungsbedarf seltener der Fall sein dürfte.

Bringt man dies mit der FaST-Konzeption in Verbindung, ist die Annahme, dass C-Fälle häufiger als statistisch erwartet erst nach regulärem Haftende entlassen werden (vor allem im Vergleich zu A-Fällen). Für diese Analyse müssen Personen mit kurzer Freiheitsstrafe ausgeschlossen werden, da sie nicht die Möglichkeit auf eine bedingte Entlassung haben. Tabelle 21 führt jene 236 Personen auf, die nach regulärem Haftende entlassen wurden (58 Personen) und bedingt entlassen wurden (178 Personen). Ein χ^2 -Test zeigt einen hochsignifikanten Zusammenhang zwischen der Art der Entlassung und der FaST-Kategorisierung.¹⁷ Es zeigte sich, dass signifikant mehr C- als A-Fälle nach dem regulären Haftende entlassen worden sind.

Personen, welche gemäss FaST einen erhöhten Abklärungsbedarf bezüglich G/S-Delinquenz aufwiesen (C-Fälle), erfüllten die Voraussetzungen für eine bedingte Entlassung überzufällig selten. Mit anderen Worten trifft das FaST jene Fälle, die sich während des Vollzugs nicht vorschriftsgemäss verhalten oder von denen angenommen wird, sie würden sich nach Freilassung nicht bewähren.

¹⁷ $\chi^2(2) = 15.71, p \leq .001$

Tabelle 21: Art der Entlassung aus dem Strafvollzug, differenziert nach FaST-Kategorien

		A	B	C	Gesamt
Entlassung nach regulärem Haftende (Strafmass mindestens 3 Monate)	<i>Absolute Häufigkeiten:</i>	26	18	14	58
	<i>Spaltenprozente:</i>	16.7%	38.3%	42.4%	100.0%
	<i>Residualwerte:</i>	-2.0	1.9	2.1	
Bedingte Entlassung aus dem Strafvollzug	<i>Absolute Häufigkeiten:</i>	130	29	19	178
	<i>Spaltenprozente:</i>	83.3%	61.7%	57.6%	75.4%
	<i>Residualwerte:</i>	1.1	-1.1	-1.2	
Gesamt	<i>Absolute Häufigkeiten:</i>	156	47	33	236
	<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%

$$\chi^2(2) = 15.71, p \leq .001$$

5.2.8 Steuerbares Einkommen nach FaST-Kategorien

Zur Einschätzung der sozialen Integration wurde jeweils das Einkommen nach einem, zwei, drei, vier, fünf und sechs Jahren Time-at-risk erhoben. Wie in Kapitel 5.1.4.6 erläutert, ist die Annahme, dass C-Fälle aus personen- und situationsbezogenen Gründen am meisten Mühe haben, sich nach Verbüsung einer Sanktion gesellschaftlich zu integrieren. Dies könnte sich in einer erhöhten Arbeitslosigkeitsrate und somit einem tieferen Einkommen widerspiegeln.

Da die Zeitspanne zwischen einem und sechs Jahren Time-at-risk bei den meisten Personen der Stichprobe die Jahre 2005-2011 betrifft, wurde das Steueramt des Kantons Zürich beauftragt, das Einkommen aller 733 Personen der Stichprobe für diese Jahre zu liefern. Es ist jedoch zu beachten, dass der Untersuchungszeitraum (2005-2011) nicht bei allen Personen der Stichprobe mit der Time-at-risk übereinstimmt. So gibt es beispielsweise Personen, deren Time-at-risk bereits früher oder erst später begann. Startete die Time-at-risk beispielsweise erst im Jahre 2008 konnte hier ein maximaler Untersuchungszeitraum von drei Jahren erreicht werden, indem die Steuerdaten des Jahres 2011 herangezogen wurden. Die Stichprobengrösse variiert deshalb über die verschiedenen Time-at-risks. Weiter meldete uns das Steueramt, wenn Personen den Kanton gewechselt oder die Schweiz verlassen hatten. Eine Vielzahl der Personen war jedoch unter den angegebenen Namen gar nicht auffindbar, weshalb viele Ausfälle zu konstatieren sind. Tabelle 22 zeigt für die jeweiligen Time-at-risks (1-6 Jahre), wie viele Personen insgesamt ausgeschlossen wurden bzw. in der Stichprobe verblieben sind. Personen, welche einen Ausschlussgrund gemäss Tabelle 5 aufweisen, wurden von Beginn an aus der Stichprobe ausgeschlossen, da für sie keine Time-at-risk ermittelt werden konnte.

Tabelle 22: Ausschlüsse und Stichprobengrößen für Auswertungen zu Steuerdaten

	Ausschluss gemäss Tabelle 5	Kantons- wechsel	Schweiz verlassen	Anderer Ausschluss- grund ¹⁸	Ausschlüsse gesamt	Verbleib in Stichprobe
Einkommen nach 1 Jahr Time-at-risk	74	43	9	390	516	217
Einkommen nach 2 Jahren Time-at-risk	74	50	16	319	459	274
Einkommen nach 3 Jahren Time-at-risk	74	49	18	277	418	315
Einkommen nach 4 Jahren Time-at-risk	74	52	23	267	416	317
Einkommen nach 5 Jahren Time-at-risk	74	58	27	273	432	301
Einkommen nach 6 Jahren Time-at-risk	74	59	26	337	496	237

Aus Tabelle 22 wird ersichtlich, dass zahlreiche Ausschlüsse zu konstatieren sind (jeweils mehr als 50% aller Fälle). Die Repräsentativität der Stichprobe für die Grundpopulation ist somit eingeschränkt. Um den Zusammenhang zwischen FaST und steuerbarem Einkommen zu überprüfen, werden zwei Zugänge verfolgt.

In einem ersten Schritt wird der Zusammenhang zwischen dem FaST und dem *blossen Vorhandensein eines steuerbaren Einkommens* (Einkommen vorhanden oder kein Einkommen vorhanden) untersucht. Aus einer inhaltlichen Sichtweise scheint nicht primär die Höhe eines erwirtschafteten Einkommens relevant, sondern vielmehr, dass überhaupt eine berufliche Tätigkeit vorhanden war, welche die soziale Integration einer Person unterstützt. Auch ein kleines Einkommen deutet darauf hin, dass zumindest über einen gewissen Zeitraum einer beruflichen Tätigkeit nachgegangen wurde. Erst in einem zweiten Schritt wird die *Höhe des steuerbaren Einkommens* nach FaST-Kategorien und verschiedenen Time-at-risks untersucht.

Tabelle 23 zeigt den Zusammenhang zwischen dem Vorhandensein von steuerbarem Einkommen (ja/nein) und den FaST-Kategorien nach verschiedenen Time-at-risks auf. Da die Daten gewichtet wurden, unterscheiden sich die Fallzahlen von jenen in Tabelle 22.

¹⁸ Person nicht identifizierbar, nicht auffindbar oder relevante Steuerjahre nicht erhoben.

Tabelle 23: Einkommen nach verschiedenen Time-at-risks, differenziert nach FaST-Kategorien

			A	B	C	Gesamt
1 Jahr Time-at-risk	Kein Einkommen	<i>Absolute Häufigkeiten:</i>	35	16	8	59
		<i>Spaltenprozente:</i>	19.1%	37.2%	27.6%	23.1%
		<i>Residualwerte:</i>	-1.1	1.9	0.5	
$p \leq .05$	Einkommen	<i>Absolute Häufigkeiten:</i>	148	27	21	196
		<i>Spaltenprozente:</i>	80.9%	62.8%	72.4%	76.9%
		<i>Residualwerte:</i>	0.6	-1.1	-0.3	
Gesamt		<i>Absolute Häufigkeiten:</i>	183	43	29	255
		<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%
2 Jahre Time-at-risk	Kein Einkommen	<i>Absolute Häufigkeiten:</i>	49	19	6	74
		<i>Spaltenprozente:</i>	23.6%	35.2%	19.4%	25.3%
		<i>Residualwerte:</i>				
$p = n.s.$	Einkommen	<i>Absolute Häufigkeiten:</i>	159	35	25	219
		<i>Spaltenprozente:</i>	76.4%	64.8%	80.6%	74.7%
		<i>Residualwerte:</i>				
Gesamt		<i>Absolute Häufigkeiten:</i>	208	54	31	293
		<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%
3 Jahre Time-at-risk	Kein Einkommen	<i>Absolute Häufigkeiten:</i>	56	22	7	85
		<i>Spaltenprozente:</i>	24.6%	42.3%	21.2%	27.2%
		<i>Residualwerte:</i>	-0.8	2.1	-0.7	
$p \leq .05$	Einkommen	<i>Absolute Häufigkeiten:</i>	172	30	26	228
		<i>Spaltenprozente:</i>	75.4%	57.7%	78.8%	72.8%
		<i>Residualwerte:</i>	0.5	-1.3	0.4	
Gesamt		<i>Absolute Häufigkeiten:</i>	228	52	33	313
		<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%
4 Jahre Time-at-risk	Kein Einkommen	<i>Absolute Häufigkeiten:</i>	40	15	9	64
		<i>Spaltenprozente:</i>	16.8%	27.8%	39.1%	20.3%
		<i>Residualwerte:</i>	-1.2	1.2	2.0	
$p \leq .05$	Einkommen	<i>Absolute Häufigkeiten:</i>	198	39	14	251
		<i>Spaltenprozente:</i>	83.2%	72.2%	60.9%	79.7%
		<i>Residualwerte:</i>	0.6	-0.6	-1.0	
Gesamt		<i>Absolute Häufigkeiten:</i>	238	54	23	315
		<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%
5 Jahre Time-at-risk	Kein Einkommen	<i>Absolute Häufigkeiten:</i>	49	15	2	66
		<i>Spaltenprozente:</i>	23.6%	28.8%	9.1%	23.4%
		<i>Residualwerte:</i>				
$p = n.s.$	Einkommen	<i>Absolute Häufigkeiten:</i>	159	37	20	216
		<i>Spaltenprozente:</i>	76.4%	71.2%	90.9%	76.6%
		<i>Residualwerte:</i>				
Gesamt		<i>Absolute Häufigkeiten:</i>	208	52	22	282
		<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%
6 Jahre Time-at-risk	Kein Einkommen	<i>Absolute Häufigkeiten:</i>	43	5	2	50
		<i>Spaltenprozente:</i>	24.0%	13.0%	12.5%	21.6%
		<i>Residualwerte:</i>	0.7	-1.0	-0.8	
$p = n.s.$	Einkommen	<i>Absolute Häufigkeiten:</i>	136	31	14	181
		<i>Spaltenprozente:</i>	76.0%	86.1%	87.5%	78.4%
		<i>Residualwerte:</i>				
Gesamt		<i>Absolute Häufigkeiten:</i>	179	36	16	231
		<i>Spaltenprozente:</i>	100.0%	100.0%	100.0%	100.0%

Signifikante Zusammenhänge sind nach einem, drei und vier, nicht jedoch nach zwei, fünf und sechs Jahren Time-at-risk zu konstatieren.¹⁹ Nach einem und drei Jahren Time-at-risk sind die B-Fälle signifikant häufiger einkommenslos als A- und C-Fälle. Annahmekonform ist einzig das Resultat bei einer Time-at-risk von vier Jahren: C-Fälle sind signifikant häufiger einkommenslos als A- und B-Fälle. Insgesamt muss aber die Befundlage als sehr heterogen eingestuft werden.

Tabelle 24 und Abbildung 4 zeigen den Zusammenhang zwischen der Höhe des steuerbaren Einkommens (Mediane und Mittelwerte) und den FaST-Kategorien auf. Es fällt auf, dass die Mittelwerte in allen Fällen höher ausfallen als die Medianwerte. Dies deutet auf Extremwerte in den oberen Bereichen hin (d.h. wenige Personen, die im Vergleich zu den anderen ein sehr hohes Einkommen aufweisen). Aussagekräftiger sind die Medianwerte (weniger Beeinflussung durch Extremwerte). Es zeigt sich, dass die A-Fälle von Beginn an ein höheres Einkommen aufweisen als B- und C-Fälle. Dieses Resultat unterstützt die Annahme, dass A-Fällen eine raschere Reintegration nach Verbüßung der Sanktion gelingt.

Auf den ersten Blick überraschend scheint hingegen, dass B-Fälle ein tieferes Einkommen erwirtschaften als C-Fälle. In Kapitel 5.1.4.6 wird die Annahme erläutert, dass Personen, die ein (eher schwereres) G/S-Delikt begangen haben (und somit häufig C-Fälle sind), aus personen- und situationsbezogenen Gründen mehr Schwierigkeiten haben, sich sozial zu reintegrieren als Personen, die kein solches Delikt begangen haben. Dies widerspiegelt sich jedoch nicht in der Höhe des Einkommens.

Aus der FaST-Konzeption heraus könnte man den Effekt jedoch damit erklären, dass B-Fälle häufig eine Deliktstruktur aufweisen, die eine soziale Reintegration erschwert. So sollten B-Fälle häufig ein „eingeschliffenes Denk- und Verhaltensmuster“ aufweisen (Bewährungs- und Vollzugsdienste Zürich, Fachsupport und Lernprogramme, 2010). Dabei kann beispielsweise an Personen mit chronischer Suchtproblematik gedacht werden.²⁰ Ein solches kann dann als gegeben betrachtet werden, wenn mehrere FaST Kriterien erfüllt sind.²¹ Die Relevanz der einzelnen FaST-Kriterien müsste jedoch im Rahmen weiterführender Studien untersucht werden.

¹⁹ 1 Jahr: $X^2(2) = 6.77, p \leq .05$
2 Jahre: $X^2(2) = 3.71, p = n.s.$
3 Jahre: $X^2(2) = 7.40, p \leq .05$
4 Jahre: $X^2(2) = 8.70, p \leq .05$
5 Jahre: $X^2(2) = 3.38, p = n.s.$
6 Jahre: $X^2(2) = 2.66, p = n.s.$

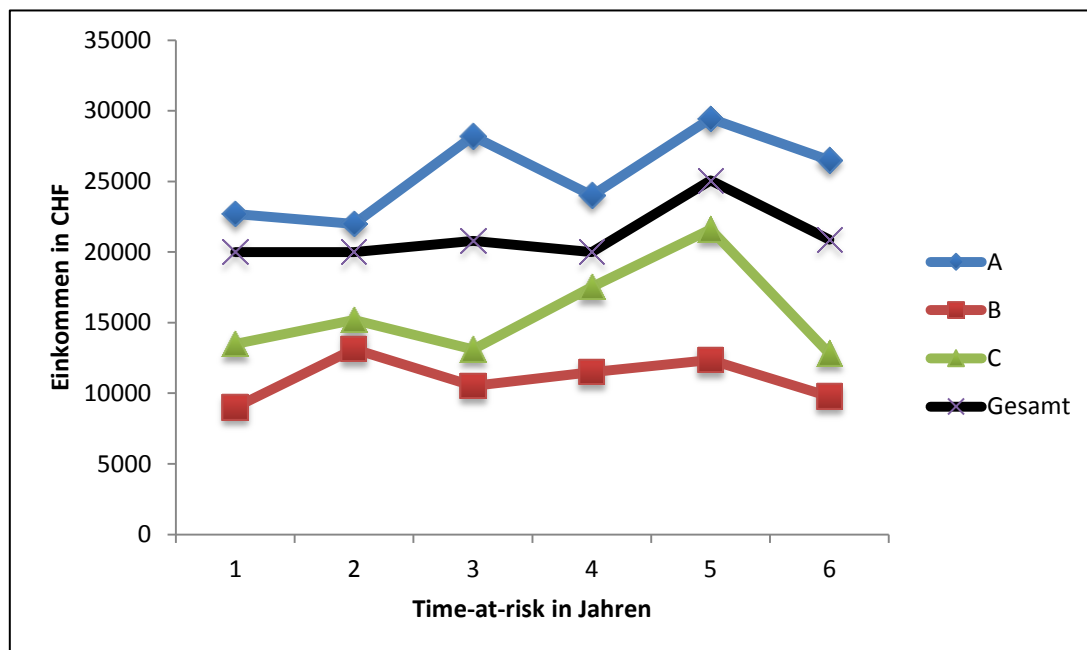
²⁰ Persönliche Kommunikation Daniel Treuthardt (12.9.2013).

²¹ Folgende Kriterien sind relevant (Bewährungs- und Vollzugsdienste, Fachsupport und Lernprogramme, 2010): „Mindestens drei Vorstrafen, mindestens drei schwerwiegende Strassenverkehrsdelikte, zwei oder mehr jugendanwaltschaftliche Einträge, Widerruf einer bedingten Strafe oder einer bedingten Entlassung, Alter bei Beginn des G/S-Deliktes < 25 Jahre, Massnahme“. Hierbei bleibt jedoch unklar, wie viele der genannten Kriterien erfüllt sein müssen.

Tabelle 24: Mittleres Einkommen in CHF, differenziert nach FaST-Kategorien und Time-at-risks

		A	B	C	Gesamt
Einkommen nach 1 Jahr Time-at-risk	<i>n:</i>	183	43	29	255
	<i>Median:</i>	22'700 CHF	9'000 CHF	13'500 CHF	20'000 CHF
	<i>Mittelwert:</i>	28'323 CHF	13'011 CHF	17'381 CHF	24'504 CHF
Einkommen nach 2 Jahren Time-at-risk	<i>n:</i>	208	54	31	293
	<i>Median:</i>	22'000 CHF	13'129 CHF	15'200 CHF	20'000 CHF
	<i>Mittelwert:</i>	29'976 CHF	17'026 CHF	23'161 CHF	26'864 CHF
Einkommen nach 3 Jahren Time-at-risk	<i>n:</i>	229	52	32	313
	<i>Median:</i>	28'208 CHF	10'518 CHF	13'100 CHF	20'797 CHF
	<i>Mittelwert:</i>	33'310 CHF	15'766 CHF	18'160 CHF	28'814 CHF
Einkommen nach 4 Jahren Time-at-risk	<i>n:</i>	238	54	23	315
	<i>Median:</i>	24'000 CHF	11'500 CHF	17'517 CHF	20'000 CHF
	<i>Mittelwert:</i>	29'754 CHF	29'423 CHF	21'911 CHF	29'124 CHF
Einkommen nach 5 Jahren Time-at-risk	<i>n:</i>	208	51	22	282
	<i>Median:</i>	29'436 CHF	12'350 CHF	21'600 CHF	25'074 CHF
	<i>Mittelwert:</i>	32'788 CHF	32'384 CHF	24'748 CHF	32'086 CHF
Einkommen nach 6 Jahren Time-at-risk	<i>n:</i>	179	36	16	231
	<i>Median:</i>	26'500 CHF	9'771 CHF	12'800 CHF	20'880 CHF
	<i>Mittelwert:</i>	32'276 CHF	27'898 CHF	18'927 CHF	20'673 CHF

Abbildung 4: Mittleres Einkommen (Mediane) in CHF, nach FaST-Kategorien und Time-at-risks



Auf varianzanalytische Auswertungen zur Bestimmung der statistischen Bedeutsamkeit der Einkommensunterschiede verschiedener FaST-Kategorien wird an dieser Stelle aufgrund der stark eingeschränkten Repräsentativität der Stichprobe verzichtet. Aus den Darstellungen in Tabelle 24 und Abbildung 4 wird jedoch ersichtlich, dass weder über die Jahre, noch über die FaST-Kategorien ein Einkommensmuster besteht.

Damit liegt der Schluss nahe, dass das Einkommen durch andere, nicht kontrollierbare Drittvariablen moderiert oder mediiert wird. Insbesondere wäre es äusserst relevant, Rückfälle zu kontrollieren. Im Rahmen der vorliegenden Studie konnte nicht erhoben werden ob, wie lange und zu welcher Zeit ein Rückfall durch eine Freiheitsstrafe verbüsst wurde. Dies wäre äusserst relevant, da bei Verbüsung einer Freiheitsstrafe nicht die Möglichkeit besteht, ein steuerbares Einkommen zu erwirtschaften. Weiter scheint relevant, ob eine Person in einer Partnerschaft lebt oder nicht. Dabei wirkt eine Partnerschaft einerseits per se als protektiver Faktor (z.B. Andrews & Bonta, 2010; Sampson, Laub & Wimer, 2006), andererseits kann sie eine „Mitfinanzierung“ durch den Partner bewirken, das heisst dazu führen, dass eine Person kein steuerbares Einkommen erwirtschaftet.

5.2.9 VRAG nach FaST-Kategorien

Um das FaST-Ergebnis der Gruppierung in A-, B- und C-Fälle anhand eines „externen“ Kriteriums validieren zu können, wurden jene Fälle der Stichprobe, wo dies möglich war, mit dem VRAG gewertet. Die konvergente Validität des FaST kann also nur *teilweise* eingeschätzt werden.

Teilweise darum, weil der VRAG sich ausschliesslich auf Fälle mit einem schweren G/S-Delikt anwenden lässt (siehe Seite 15). Eine weitere Einschränkung dieser Überprüfung ergibt sich aus dem bereits erwähnten Umstand, dass vielfach die Informationsbasis, deren die Anwendung des VRAG bedarf, nicht gegeben war, so dass der VRAG aufgrund zu vieler fehlender Angaben nicht angewendet werden konnte. Insgesamt konnte der VRAG lediglich bei 13% aller Fälle angewendet werden (siehe dazu die Ausführungen in Kapitel 5.2.9.1). Die Resultate müssen deshalb unter dem Blickwinkel betrachtet werden, dass keine trennscharfe Überprüfung der Validität der FaST-Ergebnisse vorgenommen werden konnte. Die Generalisierbarkeit der Ergebnisse auf die Gesamtstichprobe ist somit eingeschränkt. Trotzdem erlauben die Resultate eine Einschätzung darüber, ob Fälle, bei welchen ein VRAG angewendet werden konnte, auch vom FaST als abklärungsbedürftig eingestuft wurden.

Eine gute Übereinstimmung zwischen FaST und VRAG würde angezeigt, wenn der VRAG auch wirklich *überwiegend* bei den C-Fällen angewendet werden könnte, da nur diese Fälle gemäss FaST-Konzeption einer näheren Bedarfs- und Risiko-Abklärung bedürfen. Nun ist es aber so, dass die Anwendung des VRAG zwar auf G/S-Delikte beschränkt ist, dass im FaST jedoch auch Fälle mit G/S-Delikten als Anlassdelikt zu einem A- oder B-Fall werden können. Sofern der VRAG bei A- und B-Fällen angewendet werden kann, müssten diese Fälle aber *tiefe VRAG-Risikoklassen* aufweisen, da sonst das FaST diese als C-Fall hätte werten müssen.

Bei den C-Fällen kann a-priori keine bestimmte VRAG-Verteilung angenommen werden. Es ist anzunehmen, dass auch in der Kategorie der C-Fälle eine gewisse Varianz hinsichtlich der Rückfallgefahr gemäss VRAG anzutreffen ist.

Für die nachfolgenden Analysen wurden ungewichtete Daten verwendet, da nicht Aussagen zur gesamten Stichprobe gemacht werden, sondern nur zur Teilmenge der Fälle, wo ein VRAG überhaupt angewendet werden konnte. Es wurden keine Ausschlüsse vorgenommen, da sich sowohl das FaST als auch der VRAG auf das Anlassdelikt (bzw. Vorstrafen) bezieht, welches bei allen Fällen vorliegt.

5.2.9.1 Anwendungsmöglichkeiten des VRAG

Tabelle 25 zeigt, dass bei insgesamt 95 der 733 Fälle (d.h. bei 13 %) ein VRAG angewendet werden konnte. Bei 519 Fällen war kein VRAG möglich, weil kein relevantes Anlassdelikt vorlag. Bei 119 Fällen wäre die Durchführung eines VRAG aufgrund eines G/S-Delikts möglich gewesen, es lagen aber zu wenige Informationen zum Fall vor.

Bei 13 Fällen wurde der VRAG auf der Grundlage einer Vorstrafe durchgeführt. Dies war dann der Fall, wenn das Anlassdelikt beispielsweise ein Diebstahl war, die Person jedoch im Vorfeld ein G/S-Delikt begangen hatte. Dass man neben dem Anlassdelikt ebenfalls die Vorstrafen einer Person genauer betrachtet, ist ein Anliegen von ROS. Die Vorstrafen spielen bei der Bedarfs- und Risikoanalyse eine zentrale Rolle. Weiter kann man davon ausgehen, dass die VRAG-Risikoklasse über die Zeit rela-

tiv konstant bleibt, da es sich bei vielen der 12 Items um statische Faktoren handelt.²² Dass der VRAG auch aufgrund von Vorstrafen angewendet wird, stimmt mit der gängigen Praxis der BVD überein.

Tabelle 25: Anwendung des VRAG, differenziert nach FaST-Kategorie

		A	B	C	Gesamt
VRAG durchgeführt	<i>Auf Basis Anlassdelikt:</i>	14	18	50	82
	<i>Auf Basis Vorstrafe:</i>	1	4	8	13
	Gesamt:	15 (3.2%)	22 (16.9%)	58 (42.3%)	95 (13.0%)
VRAG nicht durchgeführt	<i>Kein VRAG-relevantes Delikt:</i>	401	81	37	519
	<i>Zu wenig Informationen:</i>	50	27	42	119
	Gesamt:	451 (96.8%)	108 (83.1%)	79 (57.7%)	638 (87.0%)
Gesamt		466 (100.0%)	130 (100.0%)	137 (100.0%)	733 (100.0%)

5.2.9.2 Verteilung der VRAG-Risikoklassen

Die Verteilung der VRAG-Risikoklassen ist den folgenden Darstellungen zu entnehmen (Abbildung 5, Tabelle 26).

Die mittlere VRAG-Risikoklasse beträgt 5.47 (Standardabweichung 1.61). Die VRAG-Klassen verteilen sich annähernd gleichverteilt und in keiner Weise extrem, obschon es relativ viele Werte von 7 hat. Inhaltlich ausgedrückt weisen 21 Verurteilte eine VRAG-Risikoklasse von 7 auf, was eine Rückfallwahrscheinlichkeit von 55% nach sieben Jahren bedeutet. Weitere sieben Personen weisen die Risikoklasse 8 und zwei Personen die höchste Klasse, nämlich 9 auf. Diese sind mit einem Rückfallrisiko von 76% bzw. 100% assoziiert. Insgesamt weist also knapp ein Drittel (31.6 %) der Verurteilten ein Rückfallrisiko von 55% und darüber auf.

Weitere 30% weisen eine VRAG-Risikoklasse bis zu 4 auf, was ein relativ tiefes Rückfallrisiko von 17% bedeutet. Rund 38% weisen mittlere VRAG-Risikoklassen von 5 bzw. 6 auf, was mit Rückfallrisiken zwischen 35% und 44% verbunden ist.

²² Z.B. Lebensumstände im Kindesalter des Täters, Alter des Täters zum Zeitpunkt der Deliktbegehung oder Verletzungsgrad des Opfers.

Abbildung 5: Verteilung der VRAG-Risikoklassen

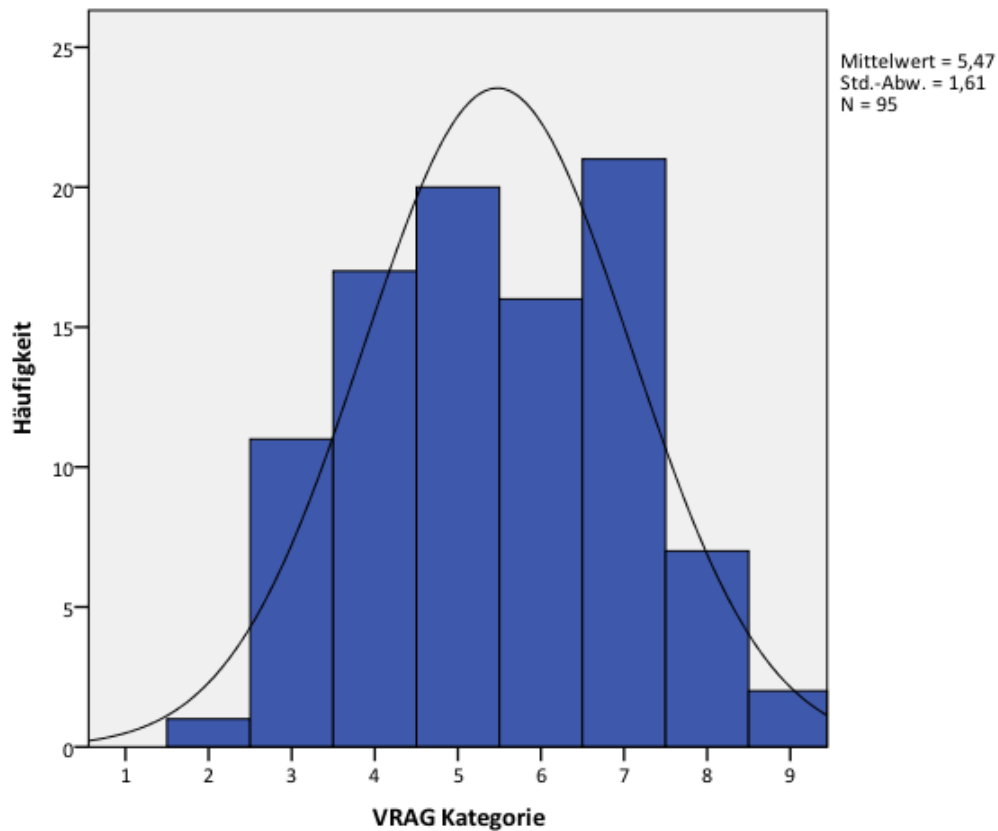


Tabelle 26: Verteilung der VRAG-Risikoklassen

VRAG-Risikoklassen und Rückfallwahrscheinlichkeiten nach 7 Jahren	Häufigkeiten	Prozente	Kumulierte Prozente
1 (0%)	0	0	0
2 (8%)	1	1.1	1.1
3 (12%)	11	11.6	12.6
4 (17%)	17	17.9	30.5
5 (35%)	20	21.1	51.6
6 (44%)	16	16.8	68.4
7 (55%)	21	22.1	90.5
8 (76%)	7	7.4	97.9
9 (100%)	2	2.1	100.0
Gesamt	95	100.0	100.0

5.2.9.3 Verteilung der VRAG-Risikoklassen nach FaST-Kategorien

Tabelle 27 zeigt die deskriptiven Werte für den VRAG separiert nach FaST-Kategorien auf. Erstens wird daraus ersichtlich, dass 58 der 95 Fälle, bei denen ein VRAG angewendet werden konnte, in die C-Kategorie fallen. 22 Fälle werden der B-, 15 Fälle gar der A-Kategorie zugeordnet. Unerwartet ist dieses Ergebnis insofern, als bei einer beachtlichen Anzahl von B- und sogar A-Fällen ein VRAG ausgefüllt werden konnte.

Zweitens wird ersichtlich, dass das durchschnittliche Rückfallrisiko gemäss VRAG nicht stetig von Kategorie A nach C zunimmt, sondern dass die B-Fälle gar einen leicht höheren VRAG-Durchschnittswert von 5.82 im Vergleich zu den C-Fällen mit 5.60 aufweisen (Differenz in Post-hoc Test Tukey nicht signifikant). Die A-Fälle weisen demgegenüber sowohl im Vergleich mit den B- als auch den C-Fällen einen signifikant tieferen Durchschnitt von 4.47 auf (Post-hoc Test Tukey beide Male signifikant).²³ Da jedoch kleine Fallzahlen vorliegen, sind die Resultate mit Vorsicht zu interpretieren. Um abschätzen zu können, ob so genannte Ausreisser (Extremwerte) die Mittelwerte verzogen haben könnten, wird zusätzlich der jeweilige Medianwert angegeben. Dieser zeigt ein ähnliches, wenn auch weniger extremes Muster auf wie die Mittelwerte.

Tabelle 27: Statistische Verteilung der VRAG-Werte, differenziert nach FaST-Kategorien

		A	B	C	Gesamt
VRAG	Anzahl Fälle:	15	22	58	95
	Mittelwert:	4.47	5.82	5.60	5.47
	Median:	5.00	6.00	5.50	5.00
	Standardabweichung:	1.36	1.47	1.64	1.61
	Minimum:	2.00	3.00	3.00	2.00
	Maximum:	7.00	9.00	9.00	9.00

Das Ergebnis wirft die Frage auf, welche Anlassdelikte zur Durchführung des VRAG führten und ob die Anlassdelikte der A- und B-Fälle als milder einzustufen sind als jene von C-Fällen. Tabelle 28 zeigt, dass die einfache Körperverletzung, insbesondere in den A- und B-Fällen, einen grossen Teil ausmacht. Dabei ist die Streuung der VRAG-Risikoklassen innerhalb der einfachen Körperverletzung breit. Man kann nicht davon ausgehen, dass Personen, welche eine einfache Körperverletzung begangen und einen A-Fall ergeben haben, in eine niedrigere VRAG-Risikoklasse eingeteilt werden als Personen, die das gleiche Delikt begangen und einen B- oder C-Fall ergeben haben.

Weiter fällt auf, dass Sexualdelikte, bei denen ein VRAG angewendet werden konnte, meist in die C-Kategorie fallen. Ist dies nicht der Fall, ergeben sich zumindest relativ tiefe VRAG-Risikoklassen von 3 (Vergewaltigung, B-Fall) oder 4 (sexuelle Nötigung, A-Fall). Auch schwere Gewaltdelikte wie Mord, Tötung, Totschlag oder schwere Körperverletzung fallen häufiger in die C-Kategorie. Wo diese in die A-Kategorie fallen, sind eher tiefe VRAG-Risikoklassen von 2-4 zu verzeichnen. Dass sich diese Fälle in der A-Kategorie befinden, ist entweder auf ein höheres Alter zum Zeitpunkt der Deliktbegehung zurückzuführen oder hat damit zu tun, dass keine weiteren FaST-relevanten Kriterien vorliegen. Konkret liegt bei den beiden A-Fällen, die einen Totschlag verübten, ein erhöhtes Alter vor (einmal 39 Jahre und einmal 65 Jahre). Jene Person, die eine schwere Körperverletzung begangen hat, wies weder Vorstrafen, noch andere relevante FaST-Kriterien auf.

²³ Gruppenunterschiede insgesamt: $F(2, 92) = 3.85, p \leq .05$

Tabelle 28: VRAG-Risikoklassen und Anlassdelikte, differenziert nach FaST-Kategorien

		VRAG-Risikoklassen									
		1	2	3	4	5	6	7	8	9	Gesamt
A	Totschlag (Art. 113 StGB)			1	1						2
	Schwere Körperverletzung (Art. 122 StGB)		1								1
	Einfache Körperverletzung (Art. 123 StGB)			3	1	4	1	1			10
	Sexuelle Nötigung (Art. 189 StGB)				1						1
	Brandstiftung (Art. 221 StGB)					1					1
	Gesamt		1	4	3	5	1	1			15
B	Einfache Körperverletzung (Art. 123 StGB)			1	2	4	2	3		1	13
	Qualifizierter Raub (Art. 140 StGB)						4	4			8
	Vergewaltigung (Art. 190 StGB)			1							1
	Gesamt			2	2	4	6	7		1	22
C	Tötung (Art. 111 StGB)				2	1					3
	Mord (Art. 112 StGB)				1						1
	Schwere Körperverletzung (Art. 122 StGB)			1				2			3
	Einfache Körperverletzung (Art. 123 StGB)				2	1	1	3	4		11
	Gefährdung des Lebens (Art. 129 StGB)				1						1
	Qualifizierter Raub (Art. 140 StGB)					1	6	1	1		9
	Freiheitsberaubung und Entführung (Art. 183 StGB)				1					1	2
	Sex. Handlungen mit Kindern (Art. 187 StGB)			2	2	4		1			9
	Sexuelle Nötigung (Art. 189 StGB)				1				1		2
	Vergewaltigung (Art. 190 StGB)			3	2	3	1	3			12
	Brandstiftung (Art. 221 StGB)							3	1		4
	Gesamt			6	12	11	8	13	7	1	58

Bei der Teilmenge der Fälle mit einem G/S-Anlassdelikt zeigt die Überprüfung mit den VRAG-Risikoklassen als externem Kriterium zur Validierung der FaST-Ergebnisse, dass die FaST-Kategorie C eine Gruppe von Verurteilten ausweist, die in Bezug auf ihr Rückfallrisiko gemäss VRAG als sehr heterogen anzusehen ist. Eine erste Folgerung daraus besteht in einer weiteren Verfeinerung des FaST-Instruments, um eine Gruppe mit substantiellem Rückfallrisiko identifizieren zu können und hernach näher abzuklären. Die ROS-Abklärung, welche mit der zweiten ROS-Etappe eingeführt wurde, nimmt eine Unterteilung der C-Fälle vor (sog. Abklärungsstufen C I bis C III).²⁴ Tabelle 29 gibt eine Übersicht über die Abklärungsstufen. Da die verschiedenen Abklärungsstufen erst mit Beginn der 2. Etappe eingeführt wurden, konnten sie für die retrospektive Studie leider nicht berücksichtigt werden. Es wäre jedoch interessant zu evaluieren, ob die drei Abklärungsformen in einem Zusammenhang mit den VRAG-Risikoklassen stehen.

Tabelle 29: Formen der Risiko-Abklärung

Abklärungsstufe	Erläuterung	Instrumente
C I	Erfolgt standardmässig bei C-Fällen wenn die Interventionsmöglichkeit unklar ist (verbleibende Zeitspanne bis Sanktionsende liegt zwischen 6 und 12 Monaten).	Kein Einsatz von Prognoseinstrumenten. Jedoch Aufzeigen des Zusammenhangs von statistisch erhärteten Risikofaktoren und dem erhöhten Delinquenzrisiko für G-/S-Delikte.
C II	Erfolgt bei C-Fällen wenn die Interventionsmöglichkeit positiv ist, jedoch kein Auftrag für eine therapeutische Intervention besteht. Wenn die vorhandenen Informationen nicht ausreichend sind wird eine Abklärung der Stufe I erstellt.	Anwendung statistischer Prognoseinstrumente und FOTRES (Forensisch Operationalisiertes Risiko-Therapie-Evaluations-System; Urbaniok, 2007)
C III	Erfolgt standardmässig bei C-Fällen wenn ein Auftrag für eine therapeutische Intervention besteht. Wenn die vorhandenen Informationen nicht ausreichend sind wird eine Abklärung der Stufe I erstellt.	Anwendung statistischer Prognoseinstrumente und FOTRES
Risikosprechstunde	Kann über die Bauchweh-Schleife bei sämtlichen Fällen erfolgen (A, B und C).	

Quelle: Bewährungs- und Vollzugsdienste Zürich, Projekt- und Qualitätsmanagement (2012).

²⁴ In der aktuellen ROS-Praxis spricht man nicht mehr von den Abklärungsstufen von C I bis C III, sondern nur noch von den Abklärungsstufen I bis III. Dies aus dem Grund, dass Abklärungen aller Stufen nicht auf C-Fälle beschränkt sind, sondern dass diese über den Weg der Risikosprechstunde auch bei A- und B-Fällen stattfinden können.

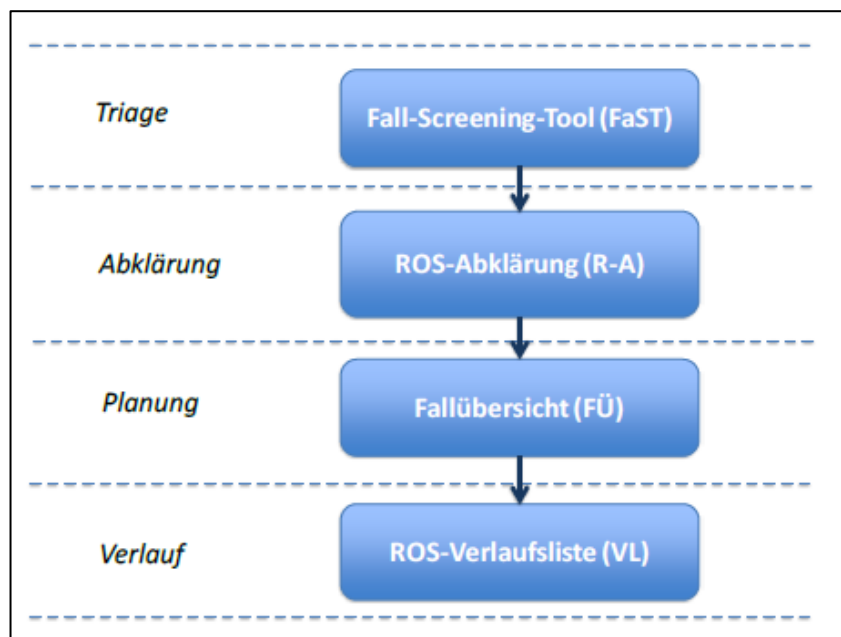
6 Evaluation von ROS-Verlaufslisten

6.1 Ziel und Hintergrund der Verlaufslisten

Im Laufe des Modellversuchs wurde die Notwendigkeit eines standardisierten Instrumentariums zur *Überprüfung des Therapieverlaufs* von Straftätern erkannt (Amt für Justizvollzug Kanton Zürich, 2012). Als solches Instrumentarium wurden sechs sog. ROS-Verlaufslisten (VL), eine Art von Checklisten, ausgearbeitet. Es sind dies die VL „Therapie“, „Alarm“, „Vollzug“, „Bewährungshilfe“, „Interventionen“ und „Motivation“. Bei den VL handelt es sich um computergestützte Tools. Das *Entwicklungsteam* der VL besteht aus Mitarbeitenden der ROS-Projektleitung und einer Fachperson für den IT-Bereich seitens BVD sowie der Abteilung Evaluation und Qualitätssicherung seitens PPD des Amtes für Justizvollzug des Kantons Zürich.

Um die VL in den Vollzugsprozess einordnen zu können, wird nachfolgend erneut das ROS-Prozessmodell dargestellt (für eine Beschreibung des gesamten ROS-Prozessmodells siehe Kapitel 1). Im Anschluss an die Phasen von Triage, Abklärung und Planung sind die Verlaufslisten (wie der Name nahelegt) für die Überprüfung des Verlaufs des weiteren Vollzugsprozesses vorgesehen.

Abbildung 6: ROS-Prozessmodell. Aus: Amt für Justizvollzug Kanton Zürich (2012)



Die Therapie soll, wie erwähnt, an den Inhalten der ROS-Abklärung ausgerichtet werden. Dies soll im Rahmen der regelmässigen Therapieberichte zum Ausdruck kommen.

Die VL erleichtern es dem Fallverantwortlichen zu prüfen, ob der Therapiebericht Bezug auf die ROS-Abklärung nimmt und inwiefern Veränderungen in den definierten Problembereichen stattgefunden haben. Dadurch wird der Fallverantwortliche unterstützt, sich in die ROS-Abklärung und den Therapiebericht einzuarbeiten und allfällig fehlende Informationen zu ermitteln. Damit kommt den VL die Funktion einer Qualitätssicherung sowohl in Bezug auf die Therapieberichte, als auch auf die Fallführung zu.

Die Anwendung der VL erfolgte in der sog. *ROS-Fokusgruppe*. Diese bestand aus 25 Mitarbeitenden aus den vier Modellversuchskantonen und wurde mit dem Ziel gebildet, die ROS-Abklärung, Fallübersicht und VL während 10 Monaten zu testen. Die Mitglieder der Fokusgruppe wurden für die Anwendung der VL geschult und waren eng an ihrem Entwicklungsprozess beteiligt. So fand ein regelmässiger Austausch zwischen dem ROS-Entwicklungsteam und den Mitgliedern der Fokusgruppe statt (zur Entwicklung der VL siehe Kapitel 6.2).

Da es sich bei den VL um ein neues Instrumentarium handelt, ist die wissenschaftliche Überprüfung der Güte eine zentrale Voraussetzung für seinen Einsatz in der Praxis. Die Entwicklung eines neuen Instruments ist durch viele Evaluations- und Anpassungsschlaufen geprägt und somit mit einem grossen zeitlichen Aufwand verbunden. Unser Auftrag bestand darin, die VL „Therapie“, welche mit einem Teil der VL „Alarm“ verknüpft ist, zu evaluieren. An dieser Stelle ist darauf hinzuweisen, dass die anderen vier VL ebenfalls einer wissenschaftlichen Evaluation bedürfen. Die BVD rechnen jedoch für die restlichen vier VL mit einem längeren Entwicklungs-, Erprobungs- und Evaluationsprozess, welcher die Modellversuchsphase überschreitet (Amt für Justizvollzug Kanton Zürich, 2012).

Die Evaluation der VL „Therapie“ besteht aus zwei Teilen: Der Überprüfung der Interraterreliabilität und konkurrenten Validität. Der Evaluation durch das Kriminologische Institut ist eine Entwicklungsphase der VL „Therapie“ vorausgegangen, die nachfolgend beschrieben wird.

6.2 Entwicklungsphase der Verlaufsliste „Therapie“

Die Entwicklungsphase der VL kann in vier Etappen eingeteilt werden (basierend auf Amt für Justizvollzug, Psychiatrisch-Psychologischer Dienst, 2013). Das ROS-Entwicklungsteam selbst nahm nach jeder Etappe eine Interraterreliabilitätsstudien vor (sog. IR1-IR4), um den Entwicklungsfortschritt wissenschaftlich zu prüfen und weitere Optimierungen der VL „Therapie“ vorzunehmen.

Im Rahmen von IR1-IR4 werteten jeweils drei oder vier Rater 20 oder 30 Fälle. Die Rater setzten sich aus Mitarbeitenden von Vollzugsdiensten und Mitarbeitenden des PPD zusammen. An den vier Studien nahmen jeweils unterschiedliche Rater teil, die dieselben Fälle codierten. Dazu wurde den Ratern pro Fall ein anonymisierter Therapiebericht und eine anonymisierte ROS-Abklärung zur Verfügung gestellt.

In einer Kick-Off Sitzung wurden die Rater instruiert, bevor sie die 30 Fälle werteten. Dabei wurde die Anwendung der VL inhaltlich und technisch besprochen. Es wurde betont, dass es wichtig sei, sich nicht gegenseitig über die Fälle oder die Anwendung der VL auszutauschen. Die Materialien wurden den Ratern ausgehändigt. Die statistische Auswertung erfolgte durch das ROS-Entwicklungsteam.

Nachdem IR1 bei mehr als der Hälfte der Items eine ungenügende Interraterreliabilität (zu tiefer Kappa-Koeffizient) hervorbrachte, wurden unter Berücksichtigung von Rückmeldungen der Rater

Anpassungen an der VL vorgenommen. Konkret wurden einzelne Items gelöscht, umformuliert oder präzisiert. Weiter wurde für ambulante und stationäre Massnahmen je eine eigene Liste formuliert. Es wurde eine standardisierte Instruktion für die Anwendung der VL ausgearbeitet.

Eine Überprüfung dieser Änderungen erfolgte in IR2. Gut 40% der Items wies jedoch eine unbefriedigende Interraterreliabilität auf. Daraufhin wurden marginale Veränderungen in der Formulierung der Items vorgenommen. Da die Rückmeldung der Rater ergeben hatte, dass mehrdeutige Berichtsinhalte vorhanden waren und dies Schwierigkeiten bereitet hatte, wurde eine weitere Antwortkategorie hinzugefügt. Neu musste der Anwender der VL also nicht mehr entscheiden, ob eine Information vorlag oder nicht („ja“ oder „nein“), sondern er hatte die Möglichkeit, die Option „unklar“ zu wählen, falls er widersprüchliche Aussagen fand.

IR3 zeigte jedoch keine massgebliche Verbesserung der Interraterreliabilität. Das ROS-Entwicklungsteam ging nach diesem Resultat davon aus, dass eine weitere Anpassung der Inhalte der VL nicht zu einem besseren Ergebnis führen würde. Vielmehr wurde gefolgert, dass aus heterogenen Therapieberichten, wie diese in der Praxis vorliegen, keine eindeutigen Informationen gezogen werden können und die VL somit nicht übereinstimmend ausgefüllt werden kann.

Das ROS-Entwicklungsteam erstellte deshalb in Zusammenarbeit mit verschiedenen Anwendergruppen der VL (Institutionen und selbstständige Therapeuten) ein standardisiertes Therapieberichtsformular, welches eine einheitliche Berichterstattung ermöglicht. Dieses besteht aus geschlossenen und offenen Fragen (bzw. die offenen Fragen spezifizieren die Antworten auf die geschlossenen Fragen).

In IR4 wurde das standardisierte Therapieberichtsformular eingesetzt, jedoch beschränkte man sich auf die geschlossenen Fragen. Konkret erhielten die Rater von IR4 sowohl den nicht-standardisierten Therapiebericht als auch das Therapieberichtsformular, welches nur die geschlossenen Fragen beinhaltete. Die Anzahl an Items mit einer guten Interraterreliabilität stieg markant an. Für das Entwicklungsteam war somit klar, dass insbesondere die heterogenen Therapieberichte für die unzureichenden Ergebnisse von IR1-IR3 verantwortlich waren. Durch diese Erkenntnis war die VL so weit in ihrem Entwicklungsprozess, dass eine externe Evaluation angemessen erschien.

Das Kriminologische Institut nahm zwei weitere Interraterreliabilitätsstudien vor (IR5 und IR6). Nach IR5 konnten somit nochmals Optimierungen vorgenommen werden, welche in IR6 überprüft wurden. Im Rahmen von IR6 wurde zusätzlich zur Reliabilität die Validität erhoben. Bevor die einzelnen Bereiche erläutert und die Resultate präsentiert werden, wird die VL „Therapie“ in ihrem Aufbau beschrieben, wie sie für IR5 und IR6 vorlag.

6.3 Aufbau der ROS-Verlaufsliste „Therapie“

Die VL „Therapie“ ist ein Excel-basiertes Tool in Form einer Checkliste und besteht aus 29 Items (siehe Anhang C, Seite 87). Die 29 Items verteilen sich auf folgende vier Hauptblöcke, wobei Block 3 und 4 jeweils nochmals in die Teile A und B unterteilt werden.

1. Rechtliche Rahmenbedingungen der Therapie (4 Items)
2. Setting und Grundlage der Therapie (5 Items)
3. Verlauf der Therapie und Beurteilung des Therapeuten (10 Items)
 - Teil A:* Allgemein (8 Items)
 - Teil B:* Items zur Medikamenten-Compliance, nur auszufüllen, wenn eine medikamentöse Behandlung vorliegt (2 Items)
4. Hinweise auf einen kritischen Verlauf der Therapie (10 Items)
 - Teil A:* Hinweise auf kritischen Therapieverlauf, kein Eintrag in die VL „Alarm“ (4 Items)
 - Teil B:* Hinweise auf kritischen Therapieverlauf, Eintrag in die VL „Alarm“ (6 Items)

Jedes Item stellt einen Checkpunkt dar. Das heisst, jedes der Items wird dahingehend beurteilt, ob die Information vorhanden ist oder nicht (vorgegebene Antwortmöglichkeiten „ja“, „nein“ bzw. „trifft zu“, „trifft nicht zu“). Ist eine Information nicht vorhanden oder bestehen Unklarheiten, hat dies je nach Item unterschiedliche Konsequenzen zur Folge. Sind Items der ersten drei Hauptblöcke betroffen, muss der Fallverantwortliche die jeweilige Information beim Therapeuten einholen, damit die Qualität des Therapieberichts gesichert ist.

Von speziellem Interesse ist jedoch der letzte Hauptblock, welcher Hinweise auf einen kritischen Therapieverlauf abfragt. Dabei gibt es zwei Abstufungen. Eine positive Wertung eines Items von Teil A zeigt dem Fallverantwortlichen auf, dass Hinweise auf einen kritischen Therapieverlauf vorliegen. Es liegt im Ermessen des Fallverantwortlichen, wie darauf zu reagieren ist. Eine positive Wertung von mindestens einem der Items von Teil B hingegen bedeutet, dass schwerwiegendere Hinweise auf einen kritischen Therapieverlauf vorhanden sind, welche einen Eintrag in die VL „Alarm“ auslösen und mit dem Vorgesetzten diskutiert werden müssen.

6.4 Interraterreliabilität der ROS-Verlaufsliste „Therapie“

IR5 und IR6 sind bezüglich der Versuchsanordnung sehr ähnlich, weshalb sie im nachfolgenden Methodenteil zusammengefasst werden. Wo Unterschiede bestehen, wird darauf hingewiesen.

6.4.1 Methode

6.4.1.1 Rater

In IR5 und IR6 werteten jeweils vier Fallverantwortliche der Fokusgruppe die Fälle. Diese Rater wurden in der Fokusgruppe für die Anwendung der VL geschult. Fallverantwortliche können entweder einen sozialarbeiterischen, juristischen oder kaufmännischen beruflichen Hintergrund aufweisen. Eine Vertretung möglichst aller Berufsgruppen ist deshalb wünschenswert. Hinsichtlich kantonaler Zugehörigkeit und beruflichem Hintergrund setzten sich die Rater wie folgt zusammen:

IR5: Alle Rater waren aus dem Kanton Zürich. Zwei Rater wiesen einen sozialarbeiterischen, einer einen juristischen und ein weiterer einen kaufmännischen beruflichen Hintergrund auf.

IR6: Zwei Rater waren aus dem Kanton Zürich, die anderen beiden aus den Kantonen Luzern bzw. St. Gallen. Je zwei Rater wiesen einen sozialarbeiterischen bzw. juristischen beruflichen Hintergrund auf.

Leider war es im Rahmen von IR6 nicht möglich, die Berufsgruppe des kaufmännischen Mitarbeitenden einzubeziehen, da in der ROS-Fokusgruppe keine weiteren Personen aus dieser Berufsgruppe zur Verfügung standen (diese wurden bereits in früheren Interraterreliabilitätsstudien eingesetzt). Da es sich aber hierbei um die seltenste Berufsgruppe bei den Fallverantwortlichen handelt, bedeutet dies keine massgebliche Einschränkung der Berufs-Repräsentativität.

6.4.1.2 Fälle

Die Rater bearbeiteten 30 anonymisierte Fälle des PPD, welche eine ambulante Therapie besuchten (gleiche Fälle für IR5 und IR6). Auf eine Überprüfung der VL für stationäre Therapien wurde aus zeitlichen Gründen verzichtet. Diese unterscheidet sich jedoch lediglich minimal von der VL für ambulante Therapien (die VL für stationäre Therapien enthält wenige zusätzliche Items). Von den 30 Fällen waren, ohne das Wissen der Rater, 15 der Therapieberichte mit Ereignissen präpariert, welche auf einen kritischen Therapieverlauf hinweisen. Diese Ereignisse führen bei mindestens einem der Items des Hauptblocks 4, Teil B, der VL „Therapie“ zu einer positiven Wertung, was wiederum einen Eintrag in die sog. „Alarmliste“ auslöst. In der Praxis müssten demnach weitere Abklärungen zusammen mit dem Vorgesetzten vorgenommen werden. Das Präparieren der Fälle wurde durch das ROS-Entwicklungsteam vorgenommen und durch dieses direkt in die Form des standardisierten Therapieberichtformulars übertragen. Ziel des Präparierens war es, für die Validitätsprüfung genügend Fälle mit Hinweisen auf einen kritischen Therapieverlauf zu haben.

6.4.1.3 Material

Zur Wertung der 30 Fälle stand den Ratern folgendes Material zur Verfügung:

- Pro Fall ein anonymisierter Therapiebericht in Form des standardisierten Therapiebericht-formulars, darin integriert eine anonymisierte Fallübersicht
- Eine Kommentarliste, in welche die Rater Bemerkungen zu einzelnen Items anbringen konnten (z.B. Anregungen für andere Formulierung, Unklarheiten bei der Interpretation eines Items etc.)
- Ein Merkblatt, auf welchem die wichtigsten Anweisungen für die Durchführung zusammengefasst wurden (siehe Anhang D, Seite 90)

Die VL „Therapie“ wurde den Ratern in elektronischer Form (Excel-Datei) zur Verfügung gestellt. Pro Fall wurde eine Excel-Datei abgespeichert.

6.4.1.4 Durchführung

Für die Übergabe der Fälle fand eine Kick-Off Sitzung mit den vier Ratern statt. Geleitet wurde sie von einer Mitarbeiterin des Kriminologischen Instituts und einem (bzw. bei IR5 zwei) Vertreter(n) des ROS-Entwicklungsteams. Es wurden folgende Punkte besprochen:

- Anwendung des standardisierten Therapieberichtformulars (war den Personen der Fokusgruppe in dieser Form noch nicht bekannt)
- Aushändigung des Merkblatts und Besprechung desselben

Sämtliches Material, mit Ausnahme der 30 Therapieberichte, wurde den Ratern nach Abschluss der Sitzung in elektronischer Form zugestellt. Die 30 Therapieberichte wurden den Ratern auf Papier ausgehändigt.

Nach Abschluss der Interraterreliabilitätsstudien liessen die Rater die ausgefüllten VL der ROS-Projektleitung in elektronischer Form zukommen, die ausgedruckten Berichte übergaben sie ebenfalls der ROS-Projektleitung.

6.4.1.5 Auswertung

Die Daten der ausgefüllten VL der vier Rater wurden durch das ROS-Entwicklungsteam aufbereitet und dem Kriminologischen Institut zur Auswertung der Interraterreliabilität übergeben. Um Rückschlüsse auf einzelne Rater zu vermeiden, nahm das Kriminologische Institut eine Anonymisierung der Rater vor.

Zur Vergleichbarkeit mit IR1-IR4 aus der Entwicklungsphase wurde für die Berechnungen der Interraterreliabilität dasselbe Reliabilitätsmass, der sog. Kappa-Koeffizient, gewählt. Der Kappa-Koeffizient kann eine Ausprägung von -1 bis +1 annehmen, wobei 1 eine völlige Übereinstimmung bedeutet und 0 eine zufällige Übereinstimmung. Negative Werte bedeuten, dass die beobachtete Übereinstimmung unterhalb der Zufallserwartung liegt, das heisst beispielsweise ein systematischer Fehler vorliegt (Gwet, 2010). Werte ab 0.6 bedeuten eine gute Übereinstimmung, während Werte zwischen 0.4 und 0.6 als befriedigend zu interpretieren sind. Liegt der Kappa-Koeffizient unter 0.4, bedeutet dies eine unbefriedigende Übereinstimmung. Beim Kappa-Koeffizienten handelt es sich um ein strenges Mass, welches sich aus dem Verhältnis zwischen der beobachteten und der erwarteten Übereinstimmung der Rater ergibt. Die erwartete Übereinstimmung ist jene, welche beim Vorliegen von reinem Zufall zustande käme.

Exkurs zum Kappa-Koeffizient

Nun wird der Kappa-Koeffizient jedoch ebenfalls durch die sog. „Prävalenz“ beeinflusst (Powers, 2012; Sim & Wright, 2005). Eine hohe Prävalenz liegt dann vor, wenn sich die Anzahl positiver (hier: JA-Antworten) und negativer Wertungen (hier: NEIN-Antworten) über die Fälle betrachtet stark voneinander unterscheidet. Folgendes Beispiel soll dies verdeutlichen:

Tabelle 30: Beispiele zur Ausprägung des Kappa-Koeffizienten, abhängig von der Prävalenz

Beispiel 1, hohe Prävalenz

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	2	2	2	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1

k = 0.619

1 = positive Wertung

2 = negative Wertung

Beispiel 2, tiefe Prävalenz

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	2	2	2	1
3	1	1	1	1
4	2	2	2	2
5	1	1	1	1
6	2	2	2	2

k = 0.832

1 = positive Wertung

2 = negative Wertung

Beispiel 1 und 2 zeigen sechs hypothetische Fälle auf, die jeweils von vier Ratern positiv oder negativ gewertet werden. Die Interraterreliabilität erfasst, inwiefern die Rater innerhalb eines Falles miteinander übereinstimmen. Es wird ersichtlich, dass bei beiden Beispielen lediglich bei Fall 2 eine Abweichung zwischen den Ratern zu konstatieren ist (siehe grüne Hinterlegung: Rater 4 wertet Fall 2 positiv, alle anderen werten ihn jedoch negativ). Bei allen anderen Fällen ist die Übereinstimmung jedoch vollständig gegeben. Die Beispiele unterscheiden sich jedoch dahingehend, dass bei Beispiel 1 fünf der sechs Fälle positiv gewertet werden (hohe Prävalenz) und bei Beispiel 2 drei der sechs Fälle positiv und die anderen drei der sechs Fälle negativ gewertet werden (tiefe Prävalenz), was zu unterschiedlichen Kappa-Werten führt. Beispiel 1 zeigt mit $k = 0.619$ zwar eine gute Übereinstimmung, die sich aber dennoch deutlich von jener von Beispiel 2 mit $k = 0.832$ unterscheidet.

Es kann gefolgert werden, dass der Kappa-Koeffizient bei hoher Prävalenz noch strenger ist als bei tiefer Prävalenz. Hohe Prävalenz liegt vermutlich bei den meisten Items der VL „Therapie“ vor, denn idealerweise, und in den meisten Fällen, ist wohl die Information im Therapiebericht vorhanden, so dass eine positive Antwort gewertet werden kann.

Für die Analysen wurde einerseits jeweils für jedes Item ein Kappa-Koeffizient berechnet, um gezielt problematische Items aufzudecken. Da die Erkennung von Hinweisen auf kritische Therapieverläufe ein zentrales Anliegen der VL ist, wurde andererseits ebenfalls mittels Kappa-Koeffizienten berechnet, inwiefern diese übereinstimmend als solche erkannt wurden.

Nach der Auswertung der Daten von IR5 und IR6 durch das Kriminologische Institut fand jeweils eine Sitzung statt, in welcher die Resultate und sich daraus ergebende Anpassungen der VL zusammen mit dem ROS-Entwicklungsteam diskutiert wurden. Die Anpassungen beinhalteten sowohl formelle, als auch inhaltliche Aspekte.

6.4.2 Ergebnisse IR5

6.4.2.1 Übersicht

Tabelle 31 zeigt für IR5 die Kappa-Koeffizienten der 29 Items der VL Therapie, basierend auf vier Ratern und 30 Fällen.

Tabelle 31: Items der VL „Therapie“ und die zugehörigen Kappa-Koeffizienten, IR5

Hinweis: BF = Standardisiertes Therapieberichtsformular; LERF = Labile eigenständig risikorelevante Faktoren

Item		Kappa
Item1	nennt das Datum des strafrechtlichen Entscheids, der zum Vollzug führte. (Siehe BF A1.1)	-0.017 unbefriedigend
Item2	nennt den gesetzlichen Auftrag (Gesetzesartikel) und gibt Auskunft über das Vorliegen eines vorzeitigen Massnahmenantritts bzw. einer freiwilligen Therapie. (Siehe BF A1.3;1.4)	0.322 unbefriedigend
Item3	nennt die gerichtlich und/oder behördlich angeordneten Weisungen / Auflagen. (Wenn keine angeordnet "ja" ankreuzen) (Siehe BF A1.5)	0.506 befriedigend
Item4	nennt das Datum, an dem der Auftrag zur Therapie erteilt wurde (Tag, Monat, Jahr). (Siehe BF A: Datum der Therapievereinbarung bzw. des Vollzugsauftrags)	1.000 gut
Item5	nennt das Datum des Therapiebeginns (Monat, Jahr). (Siehe BF B1.1)	1.000 gut
Item6	nennt den Zeitraum, auf den sich der Therapiebericht bezieht (Nennung von Anfang und Ende [Monat, Jahr]). (Siehe BF B1.2)	1.000 gut
Item7	nennt die Frequenz der Therapiesitzungen im Berichtszeitraum. (Siehe BF B1.3)	-0.008 unbefriedigend
Item8	nennt das Gesprächssetting, in welchem die Therapiesitzungen stattgefunden haben: Einzel- und/oder Gruppentherapie. (Siehe BF B1.6)	1.000 gut
Item9	nennt die ROS-Abklärung als Informationsgrundlage für den Bericht. (Siehe BF B2.1)	1.000 gut
Item10	stellt dar, ob der Therapeut beim Klienten eine Diagnose nach ICD / DSM als vorliegend erachtet. (Siehe BF C1.2)	-0.008 unbefriedigend
Item11	beschreibt, dass alle im Fallkonzept (ROS-Abklärung) dargestellten Problembereiche für die Therapie übernommen wurden. (Siehe BF C2.2)	0.924 gut
Item12	enthält Aussagen über die Zuverlässigkeit des Klienten beim Einhalten der Termine. (Siehe BF D4.1)	-0.089 unbefriedigend
Item13	enthält Angaben über die Einhaltung von gerichtlich und/oder behördlich angeordneten Weisungen und/oder Auflagen. (Wenn keine angeordnet "ja" ankreuzen) (Siehe BF D5.1)	0.202 unbefriedigend
Item14	enthält Aussagen über die in der Fallübersicht definierten deliktrelevanten äusseren Faktoren (LERF). (Siehe BF D7.1)	-0.017 unbefriedigend
Item15	beschreibt die Zweckmässigkeit der aktuellen Rahmenbedingungen der Therapie (Therapiesetting, -frequenz). (Siehe BF D9.1)	-0.304 unbefriedigend
Item16	weist das Rückfallrisiko des Klienten zu Beginn der Therapie gemäss (mind.) eines der nachfolgenden mechanischen Risk-Assessment Instruments aus: FOTRES, VRAG, SORAG, Static-99, Static-2002, ODARA, DVRAG. (Siehe BF E1)	1.000 gut
Item17	beschreibt die Einschätzung des Therapeuten bezüglich des Rückfallrisikos zum Berichtszeitpunkt. (Siehe BF E2-4)	-0.228 unbefriedigend

Item		Kappa
Item18	enthält Aussagen über die Medikamenten-Compliance des Klienten. (Siehe BF D6.1a)	0.857 gut
Item19	enthält Aussagen über die Ansprechbarkeit des Klienten auf die Medikamente. (Siehe BF D6.1b)	0.857 gut
Item20	verdeutlicht, dass weitere deliktrelevante Problembereiche erkannt wurden, die nicht im Fallkonzept (ROS-Abklärung) festgehalten sind. (Siehe BF C2.2)	0.799 gut
Item21	weist aus, dass im Berichtszeitraum keine deliktpräventive Veränderung in den Problembereichen gemäss Fallkonzept (ROS-Abklärung) beobachtbar war. (Siehe BF D3.1)	0.718 gut
Item22	weist eine ungenügende Zuverlässigkeit des Klienten beim Einhalten von Terminen aus. (Siehe BF D4.1)	-0.333 unbefriedigend
Item23	weist aus, dass die aktuellen Rahmenbedingungen der Therapie (Therapie-setting, -frequenz) nicht oder nur fraglich geeignet sind, um dem Rückfallrisiko zweckmässig zu begegnen. (Siehe BF D9.1)	1.000 gut
Item24	verdeutlicht, dass nicht alle deliktrelevanten Problembereiche gemäss ROS-Fallkonzept in den Fokus der Therapie gerückt sind. (Siehe BF D1.2)	0.511 befriedigend
Item25	verdeutlicht, dass der Therapeut an keinem der deliktrelevanten Problembereiche gemäss Fallkonzept (ROS-Abklärung) gearbeitet hat. (Siehe BF D1.3)	0.569 befriedigend
Item26	weist aus, dass gerichtlich und/oder behördlich angeordnete Weisungen und/oder Auflagen nicht eingehalten wurden. (Wenn keine angeordnet "trifft nicht zu" ankreuzen) (Siehe BF D5.1)	0.480 befriedigend
Item27	weist eine ungenügende Compliance des Klienten in Bezug auf die Einnahme der zur Rückfallprävention wichtigen Medikamente aus. (Siehe BF D6.1a)	0.557 befriedigend
Item28	weist auf eine kritische Entwicklung in den in der Fallübersicht definierten deliktrelevanten äusseren Faktoren (LERF) hin. (Siehe BF D7.3)	-0.017 unbefriedigend
Item29	verdeutlicht, dass die in der Fallübersicht definierten deliktrelevanten äusseren Faktoren (LERF) nicht hinreichend überprüft wurden. (Siehe BF D7.3, wenn "unbekannt" oder nicht ausgefüllt)	-0.020 unbefriedigend

gut: $k \geq 0.600$

befriedigend: $0.400 \leq k < 0.600$

unbefriedigend: $k < 0.400$

Tabelle 32 gibt die Anteile an Items mit guter, befriedigender und unbefriedigender Übereinstimmung an. Zusätzlich zu IR5 sind die Übereinstimmungen von IR1 bis IR4 angegeben, um den Entwicklungsprozess sichtbar zu machen. Ein direkter Vergleich der einzelnen Interraterreliabilitätsstudien ist jedoch nur bedingt möglich, da die Fälle nicht in allen Studien dieselben waren. Veränderungen der Interraterreliabilität könnten somit auch auf die wechselnden Fälle (und nicht bloss auf die VL) zurückzuführen sein. In IR5 zeigten jeweils gut 40% der Items eine gute bzw. unbefriedigende Übereinstimmung. Bei rund 17% der Items ist eine befriedigende Übereinstimmung zu konstatieren.

Tabelle 32: Zusammenfassung der Übereinstimmung, IR1 bis IR5

	IR1	IR2	IR3	IR4	IR5
Gute Übereinstimmung	26.6%	26.1%	26.7%	48.0%	41.4%
Befriedigende Übereinstimmung	16.7%	16.7%	16.7%	0.0%	17.2%
Unbefriedigende Übereinstimmung	56.7%	41.7%	44.0%	52.0%	41.4%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

In Detailanalysen wurden die Items mit unbefriedigender Interraterreliabilität analysiert und zusammen mit dem ROS-Entwicklungsteam diskutiert (die Detailanalysen sämtlicher Items finden sich in Anhang E, Seite 91). Dabei wurden die Kommentarlisten, in welchen die Rater Unklarheiten und Unstimmigkeiten mitteilen konnten, beigezogen. Aus den Resultaten und den Kommentarlisten abgeleitet wurden folgende Anpassungen vorgenommen:

- Instruktionen für die Anwendung der VL wurden spezifiziert oder intuitiver dargestellt
- Anpassungen bei den Formulierungen einzelner Items der VL
- Anpassungen des Therapieberichtformulars, damit eine bessere Übereinstimmung mit der VL gewährleistet werden kann

6.4.2.2 Interraterreliabilität bezüglich Hinweisen auf einen kritischen Therapieverlauf IR5

Die Items 1 bis 19 tragen zur Qualitätssicherung bei. Die Items 20 bis 29 sollen es ermöglichen, wie auf Seite 51 erläutert, Hinweise auf kritische Therapieverläufe in zwei Abstufungen wahrzunehmen (Items 20-29 und 24-29). Nun ist es zentral zu wissen, ob Hinweise auf einen kritischen Therapieverlauf einstimmig von allen Ratern als vorliegend erachtet wurden. Um dies zu prüfen, wurden zwei neue Variablen gebildet. Die erste zeigt auf, ob bei der Wertung der Items 20-29 mindestens ein Hinweis auf einen kritischen Therapieverlauf vorhanden ist (d.h. mindestens eines der Items 20-29 positiv gewertet wurde). Die zweite Variable zeigt auf, ob bei den Items 24-29 mindestens eine positive Wertung vorlag.

Tabelle 33 zeigt die beiden neu berechneten Variablen im Detail auf. Die Interraterreliabilität über alle 10 Items, die auf einen kritischen Therapieverlauf hinweisen, liegt bei $k = 0.372$, was eine unbefriedigende Übereinstimmung bedeutet. Die Interraterreliabilität der sechs Items, die einen Eintrag in die Alarmliste auslösen weist mit $k = 0.518$ eine befriedigende Übereinstimmung auf.

Tabelle 33: Erkennung von Hinweisen auf einen kritischen Therapieverlauf, IR5²⁵

Hinweise auf einen kritischen Therapieverlauf
(Items 20-29)

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	2	1
2	2	2	2	2
3	2	2	2	2
4	1	1	2	1
5	2	2	2	2
6	2	2	2	1
7	2	2	2	2
8	2	2	2	2
9	1	1	2	1
10	1	1	2	1
11	1	1	2	1
12	1	1	2	1
13	1	1	2	1
14	2	2	2	2
15	2	2	2	2
16	1	1	2	1
17	2	2	2	2
18	1	1	2	1
19	1	1	2	1
20	2	1	2	1
21	1	1	2	1
22	2	1	2	2
23	2	2	2	2
24	1	1	2	1
25	1	1	2	1
26	2	1	2	1
27	2	2	2	2
28	2	2	2	2
29	2	2	2	2
30	2	2	2	2

k = 0.372

1 = kein Hinweis auf kritischen Verlauf
2 = mind. 1 Hinweis auf kritischen Verlauf

Hinweise auf einen kritischen Therapieverlauf
(Items 24-29)

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	2	2	1
3	2	2	2	2
4	1	1	1	1
5	1	1	1	2
6	2	1	2	1
7	2	2	2	2
8	1	1	2	1
9	1	1	2	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	1	2	1
15	1	2	2	1
16	1	1	1	1
17	1	1	2	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	2	2	2	2
24	1	1	1	1
25	1	1	2	1
26	1	1	1	1
27	2	2	2	2
28	1	1	2	1
29	2	1	2	2
30	2	2	2	2

k = 0.518

1 = kein Hinweis auf kritischen Verlauf
2 = mind. 1 Hinweis auf kritischen Verlauf

Dies bedeutet, dass Hinweise auf einen kritischen Therapieverlauf im Rahmen von IR5 nur teilweise befriedigend übereinstimmend erkannt wurden.

Die gemeinsam diskutierten formellen und inhaltlichen Anpassungen im Hinblick auf IR6 sollten sich nicht nur auf Item-Ebene, sondern auch auf die übereinstimmende Erkennung von kritischen Therapieverläufen positiv auswirken.

²⁵ Bei dieser Art von Tabellen zeigen die farbigen Hinterlegungen einzelner Zellen auf, wo jeweils Diskrepanzen in den Wertungen zwischen den Ratern bestehen. Bei ungenügender Interraterreliabilität sind die Hinterlegungen rot, bei genügender Interraterreliabilität blau und bei guter Interraterreliabilität grün eingefärbt.

6.4.3 Ergebnisse IR6

6.4.3.1 Übersicht IR6

Tabelle 34 zeigt für IR6 die Kappa-Koeffizienten der 29 Items der VL „Therapie“, basierend auf vier Ratern und 30 Fällen.

Tabelle 34: Items der VL „Therapie“ und die zugehörigen Kappa-Koeffizienten, IR6

Hinweis: BF = Standardisiertes Therapieberichtsformular; LERF = Labile eigenständig risikorelevante Faktoren

Item		Kappa
Item1	nennt das Datum, an dem der Auftrag zur Therapie erteilt wurde (Tag, Monat, Jahr). (Siehe BF A1.1)	1.000 gut
Item2	nennt das Datum des strafrechtlichen Entscheids, der zum Vollzug führte. (Siehe BF A2.1)	1.000 gut
Item3	nennt den gesetzlichen Auftrag (Gesetzesartikel) und gibt Auskunft über das Vorliegen eines vorzeitigen Massnahmenantritts bzw. einer freiwilligen Therapie. (Siehe BF A2.3)	-0.001 unbefriedigend
Item4	nennt die gerichtlich und/oder behördlich angeordneten Weisungen/Auflagen. (Wenn keine angeordnet, ""ja"" ankreuzen) (Siehe BF A2.4)	0.869 gut
Item5	nennt die ROS-Abklärung als Informationsgrundlage für den Bericht. (Siehe BF B1.1)	0.840 gut
Item6	nennt das Datum des Therapiebeginns (Monat, Jahr). (Siehe BF B3.1)	1.000 gut
Item7	nennt den Zeitraum, auf den sich der Therapiebericht bezieht (Nennung von Anfang und Ende [Monat, Jahr]). (Siehe BF B3.2)	1.000 gut
Item8	nennt die Anzahl der Therapiesitzungen im Berichtszeitraum. (Siehe BF B3.3)	1.000 gut
Item9	nennt das Gesprächssetting, in welchem die Therapiesitzungen stattgefunden haben: Einzel- und/oder Gruppentherapie. (Siehe BF B3.3)	1.000 gut
Item10	stellt dar, ob der Therapeut beim Klienten eine Diagnose nach ICD / DSM als vorliegend erachtet. (Siehe BF C1.2)	-0.053 unbefriedigend
Item11	beschreibt, dass alle im Fallkonzept (ROS-Abklärung) dargestellten Problembereiche für die Therapie übernommen wurden. (Siehe BF C2.3)	1.000 gut
Item12	enthält Aussagen über die Zuverlässigkeit des Klienten beim Einhalten der Termine. (Siehe BF D2.2)	1.000 gut
Item13	enthält Angaben über die Einhaltung von gerichtlich und/oder behördlich angeordneten Weisungen und/oder Auflagen. (Wenn keine angeordnet "ja" ankreuzen) (Siehe BF D2.3)	1.000 gut
Item14	enthält Aussagen über die im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF). (Siehe BF D4.1)	1.000 gut
Item15	nimmt Stellung zur Zweckmässigkeit der aktuellen Rahmenbedingungen der Therapie (Therapiesetting, -frequenz). (Siehe BF D6.1)	1.000 gut
Item16	weist das Rückfallrisiko des Klienten zu Beginn der Therapie gemäss mind. eines der nachfolgenden mechanischen Risk-Assessment Instrumente aus: FOTRES, VRAG, SORAG, Static-99, Static-2002, ODARA, DVRAG. (Siehe BF E1)	1.000 gut
Item17	beschreibt die aktuelle Einschätzung des Therapeuten bezüglich des Rückfallrisikos zum Berichtszeitpunkt. (Siehe BF E3)	-0.017 unbefriedigend
Item18	enthält Aussagen über die Medikamenten-Compliance des Klienten. (Siehe BF D3.1a)	0.739 gut
Item19	enthält Aussagen über die Ansprechbarkeit des Klienten auf die Medikamente. (Siehe BF D3.1b)	0.739 gut

Item		Kappa
Item20	verdeutlicht, dass weitere deliktrelevante Problembereiche erkannt wurden, die nicht im Fallkonzept (ROS-Abklärung) festgehalten sind. (Siehe BF C2.3)	0.923 gut
Item21	weist eine ungenügende Zuverlässigkeit des Klienten beim Einhalten von Terminen aus. (Siehe BF D2.2)	1.000 gut
Item22	weist aus, dass im Berichtszeitraum keine deliktpräventive Veränderung in den Problembereichen gemäss Fallkonzept (ROS-Abklärung) beobachtbar war. (Siehe BF D5.1)	0.933 gut
Item23	weist aus, dass die aktuellen Rahmenbedingungen der Therapie (Therapiesetting, -frequenz) nicht oder nur fraglich geeignet sind, um dem Rückfallrisiko zweckmässig zu begegnen. (Siehe BF D6.1)	1.000 gut
Item24	verdeutlicht, dass nicht alle deliktrelevanten Problembereiche gemäss ROS-Fallkonzept in den Fokus der Therapie gerückt sind. (Siehe BF D1.1)	0.815 gut
Item25	verdeutlicht, dass der Therapeut an keinem der deliktrelevanten Problembereiche gemäss Fallkonzept (ROS-Abklärung) gearbeitet hat. (Siehe BF D1.2)	0.518 befriedigend
Item26	weist aus, dass gerichtlich und/oder behördlich angeordnete Weisungen und/oder Auflagen nicht eingehalten wurden. (Wenn keine angeordnet "trifft nicht zu" ankreuzen) (Siehe BF D2.3)	0.782 gut
Item27	weist eine ungenügende Compliance des Klienten in Bezug auf die Einnahme der zur Rückfallprävention wichtigen Medikamente aus. (Siehe BF D3.1a)	0.766²⁶ gut
Item28	weist auf eine kritische Entwicklung in den im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF) hin. (Siehe BF D4.2)	0.322 unbefriedigend
Item29	verdeutlicht, dass die im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF) nicht hinreichend überprüft wurden. (Siehe BF D4.2, wenn "unbekannt" oder nicht ausgefüllt, "trifft zu" ankreuzen)	0.791 gut

gut: $k \geq 0.600$

befriedigend: $0.400 \leq k < 0.600$

unbefriedigend: $k < 0.400$

Zusammengefasst ergibt sich bei 82.8% der Items eine gute Übereinstimmung, was eine Verdopplung im Vergleich zu IR5 bedeutet. Dementsprechend wiesen in IR6 auch weniger Items befriedigende und unbefriedigende Übereinstimmungen auf (siehe Tabelle 35).

Tabelle 35: Zusammenfassung der Übereinstimmung, IR1 bis IR6

	IR1	IR2	IR3	IR4	IR5	IR6
Gute Übereinstimmung	26.6%	26.1%	26.7%	48.0%	41.4%	82.8%
Befriedigende Übereinstimmung	16.7%	16.7%	16.7%	0.0%	17.2%	3.5%
Unbefriedigende Übereinstimmung	56.7%	41.7%	44.0%	52.0%	41.4%	13.8%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Für die vier Items, welche eine unbefriedigende Interraterreliabilität aufweisen, werden nachfolgend Detailanalysen dargestellt.²⁷ Mögliche Ursachen wurden zusammen mit dem ROS-Entwicklungsteam und unter Zuhilfenahme der Kommentarlisen der vier Rater diskutiert. Daraus abgeleitet werden marginale Anpassungen an den Items vorgenommen, welche nachfolgend vermerkt sind.

²⁶ Hinweis zu Item 27: Rater 3 hat Item 27 jeweils nicht ausgefüllt, wenn Items 18 und 19 (Medikamenten-Compliance) nicht ausgefüllt wurden. Item 27 wurde in diesen Fällen recodiert. Streng genommen ergäbe sich durch dieses Instruktions-Missverständnis jedoch eine ungenügende Interraterreliabilität für Item 27 ($k = 0.104$). Eine automatisierte Führung durch die Items (z.B. automatisches Ausfüllen von Item 27) ist in Abklärung.

²⁷ Zur Vollständigkeit werden die Detailanalysen sämtlicher Items in Anhang F (Seite 91) aufgeführt.

6.4.3.2 Detailanalysen auf Item-Ebene, unbefriedigende Interraterreliabilitäten IR6

Item 3

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	2	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	2	1	1
16	1	1	1	1
17	1	1	1	1
18	1	2	1	1
19	1	1	1	1
20	1	1	1	1
21	1	2	2	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	2	1	1
27	1	2	1	1
28	1	2	1	1
29	1	2	1	1
30	1	1	1	1

$k = -0.001$

1 = ja; 2 = nein

Item 3: nennt den gesetzlichen Auftrag (Gesetzesartikel) und gibt Auskunft über das Vorliegen eines vorzeitigen Massnahmenantritts bzw. einer freiwilligen Therapie. (Siehe BF A2.3).

Anpassung: Aus der Kommentarliste von Rater 2 wurde ersichtlich, dass es zu Schwierigkeiten kam, wenn das Urteil nicht rechtskräftig war. Um diesem Umstand Rechnung zu tragen, wird das Item neu folgendermassen formuliert:

„nennt den gesetzlichen Auftrag (Gesetzesartikel) **bzw.** gibt Auskunft über das Vorliegen eines vorzeitigen Massnahmenantritts **oder** einer freiwilligen Therapie. (Siehe BF A2.3).“

Item 10

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	2	1
6	1	1	1	1
7	1	1	2	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	2	1
13	1	1	1	1
14	1	1	2	1
15	1	1	1	1
16	1	1	2	2
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	2	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

$k = -0.053$

1 = ja; 2 = nein

Item 10: stellt dar, ob der Therapeut beim Klienten eine Diagnose nach ICD / DSM als vorliegend erachtet. (Siehe BF C1.2).

Anpassung: Es wurde ersichtlich, dass Rater 3 bei Fällen, wo im standardisierten Therapieberichtformular zwar eine Diagnose, jedoch kein Verweis auf ICD oder DSM vorhanden war, konsequent mit „nein“ geantwortet hat. Da die Bezugnahme auf diese psychiatrischen Krankheitsklassifikationssysteme ein wichtiges Qualitätsmerkmal ist, sollte dies in Therapieberichten immer vermerkt werden. An der VL werden keine Änderungen vorgenommen.

Item 17

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	2	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	2	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

$k = -0.017$

1 = ja; 2 = nein

Item 17: beschreibt die aktuelle Einschätzung des Therapeuten bezüglich des Rückfallrisikos zum Berichtszeitpunkt. (Siehe BF E3).

Anpassung: BF E3 beinhaltet keine Beschreibung der aktuellen Einschätzung des Rückfallrisikos, sondern vielmehr eine Benennung desselben. Das Item wird folgendermassen umformuliert:
„**benennt** die aktuelle Einschätzung des Therapeuten bezüglich des Rückfallrisikos zum Berichtszeitpunkt. (Siehe BF E3).“

Item 28

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	2	2	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

$k = 0.322$

1 = trifft nicht zu; 2 = trifft zu

Item 28: weist auf eine kritische Entwicklung in den im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF) hin. (Siehe BF D4.2).

Anpassung: Das Item beinhaltet eine indirekte Verknüpfung, nämlich muss der Anwender der VL merken, dass eine kritische Entwicklung auch vorliegt, wenn keine Angaben zu den LERF gemacht werden. Deshalb wird das Item folgendermassen spezifiziert:

„weist auf eine kritische Entwicklung in den im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF) hin bzw. macht keine Angaben dazu (Siehe BF D4.2).“

6.4.3.3 Interraterreliabilität bezüglich Hinweisen auf einen kritischen Therapieverlauf IR6

Wie bei IR5 wurde auch im Rahmen von IR6 geprüft, inwiefern ein kritischer Therapieverlauf von allen Ratern einstimmig als solcher erkannt wurde. Wiederum wurden separate Auswertungen für die Items 20-29 und 24-29 vorgenommen. Tabelle 36 zeigt die Resultate für die beiden neu berechneten Variablen im Detail auf. Die Interraterreliabilität über alle 10 Items, die auf einen kritischen Therapieverlauf hinweisen, lag bei $k = 0.822$, was eine gute Übereinstimmung bedeutet. Die Interraterreliabilität der sechs Items, die einen Eintrag in die Alarmliste auslösen, wies eine geringere, aber immer noch gute Übereinstimmung von $k = 0.675$ auf.

Tabelle 36: Erkennung von Hinweisen auf einen kritischen Therapieverlauf, IR6

Hinweise auf einen kritischen Therapieverlauf
(Items 20-29)

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	2
2	1	2	2	2
3	2	2	2	2
4	1	1	1	1
5	2	2	2	2
6	2	2	2	2
7	2	2	2	2
8	2	2	2	2
9	2	2	2	2
10	2	2	2	2
11	1	1	1	1
12	1	1	1	1
13	2	2	2	2
14	2	2	2	2
15	2	2	2	2
16	2	2	2	2
17	2	2	2	2
18	1	1	1	1
19	2	2	2	2
20	1	2	2	2
21	1	1	1	1
22	2	2	2	2
23	2	2	2	2
24	1	1	1	1
25	1	2	2	2
26	2	2	2	2
27	2	2	2	2
28	2	2	2	2
29	2	2	2	2
30	2	2	2	2

$k = 0.822$

1 = kein Hinweis auf kritischen Verlauf

2 = mind. 1 Hinweis auf kritischen Verlauf

Hinweise auf einen kritischen Therapieverlauf
(Items 24-29)

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	2
2	1	1	1	1
3	2	2	2	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	2	2	2	2
8	1	1	1	1
9	1	1	2	2
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	2	2	2
15	2	2	2	1
16	2	2	2	1
17	2	2	2	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	2	2	2	2
24	1	1	1	1
25	1	2	2	2
26	1	1	1	1
27	2	2	2	2
28	1	1	1	2
29	1	2	2	1
30	2	2	2	2

$k = 0.675$

1 = kein Hinweis auf kritischen Verlauf

2 = mind. 1 Hinweis auf kritischen Verlauf

6.5 Validität der ROS-Verlaufslisten „Therapie“ und „Alarm“

Die gute Interraterreliabilität bezüglich der Erkennung von kritischen Therapieverläufen im Rahmen von IR6 ist eine wichtige, jedoch noch keine hinreichende Voraussetzung für den Einsatz der VL „Therapie“ in der Praxis. Zusätzlich ist es wichtig, dass die „richtigen“ Fälle als problematisch eingestuft werden. So könnte es beispielsweise sein, dass alle vier Rater einen Hinweis auf einen kritischen Therapieverlauf als vorliegend erachten, obwohl keiner vorhanden ist oder umgekehrt (wobei hohe Reliabilitäten resultierten). Dies wird anhand der konkurrenten Validität, einer Form der Kriteriumsvalidität eruiert (Hussy, Schreier & Echterhoff, 2013).

6.5.1 Methode

Die Überprüfung der konkurrenten Validität wurde anhand der Daten von IR6 vorgenommen. Die 30 Fälle wurden von zwei Personen des ROS-Entwicklungsteams mit den VL gewertet, welche ihre Lösungen aufeinander abstimmten. Dadurch generierten sie gemeinsam Musterlösungen für die 30 Fälle von IR6.

Für die Auswertungen wurde jeweils jeder Rater mit der Musterlösung bezüglich der Erkennung von Hinweisen auf einen kritischen Therapieverlauf (wiederum zweistufig: Items 20-29 und Items 24-29) verglichen. Es wurden Analysen zu Kappa-Werten, prozentualen Übereinstimmungen sowie ROC-Kurven („receiver operating characteristics“; Mossman, 1994) durchgeführt.

Die ROC-Kurve wird in einem Koordinatensystem abgebildet und ist die Funktion des Verhältnisses der Richtig-Positiven zu den Falsch-Negativen. Die Fläche unter der Kurve ist die sog. AUC („area under the curve“) und gibt an, in welchem Ausmass ein Modell zwischen zwei möglichen Ergebnissen (hier: Hinweis auf einen kritischen Therapieverlauf versus kein Hinweis auf einen kritischen Therapieverlauf) diskriminieren kann. Durch einen Vergleich der Musterlösung mit den Wertungen der Rater kann auf diese Weise die konkurrente Validität der VL „Therapie“ untersucht werden.

Die AUC kann Werte von 0 bis 1 annehmen, wobei 0.5 eine zufällige Übereinstimmung bedeutet. Je grösser die AUC-Fläche (d.h. Werte > 0.7), desto besser die Validität. Hosmer und Lemeshow (2000) schlagen folgende Faustregeln vor:²⁸

AUC = 0.5:	Keine Diskriminationsfähigkeit (bzw. Zufallsresultat)
0.7 ≤ AUC < 0.8:	Akzeptable Diskriminationsfähigkeit
0.8 ≤ AUC < 0.9:	Gute Diskriminationsfähigkeit
AUC ≥ 0.9:	Sehr gute Diskriminationsfähigkeit

6.5.2 Ergebnisse

Die beiden nachfolgenden Tabellen zeigen die Kappa-Werte, die prozentualen Übereinstimmungen und die AUC-Werte mit den zugehörigen Konfidenzintervallen (KI) auf. Dabei werden in einem ersten Schritt die Werte aller vier Rater mit den Werten der Musterlösung verrechnet. In einem zweiten Schritt werden daraus Durchschnittswerte gebildet. Tabelle 37 berücksichtigt die Items 20-29, welche

²⁸ Diese können als grobe Richtlinien für die Interpretation der Diskriminationsfähigkeit angesehen werden. Swets (1988) beispielsweise schlägt eine konservativere Kategorisierung vor, wobei folgendes gilt: AUCs zwischen 0.5 und 0.7 weisen auf eine geringe Diskriminierungsfähigkeit, Werte zwischen 0.7 und 0.9 auf eine moderate und Werte von über 0.9 auf eine hohe Diskriminationsfähigkeit hin.

auf einen kritischen Therapieverlauf hinweisen. Tabelle 38 schliesst hingegen nur die Alarm-Items 24-29 ein.

Es wird ersichtlich, dass insbesondere bei Berücksichtigung der Items 20-29 eine gute Interraterreliabilität (durchschnittlicher Kappa-Wert = 0.92) und eine sehr gute Validität (durchschnittlicher AUC-Wert = 0.96) vorliegen. Bezieht man nur die Items 24-29 in die Auswertungen ein, sind die Reliabilitäts- und Validitätswerte etwas tiefer. Der Kappa-Wert von 0.60 ist jedoch immer noch als gut zu bewerten und auch der AUC-Wert von 0.80 ist als gut einzustufen.

Dass die Items 20-29 bessere Werte aufweisen als die Items 24-29 liegt unseres Erachtens daran, dass mit zunehmender Anzahl an zu beantwortenden Items die Wahrscheinlichkeit steigt, dass ein kritischer Therapieverlauf tatsächlich als solcher erkannt wird (was jedoch auch abhängig von der Anzahl Präparierungsmerkmale pro Fall ist).²⁹

Tabelle 37: Interraterreliabilität und konkurrente Validität, Items 20-29

Grundlage	Kappa	Prozentuale Übereinstimmung	AUC [95% KI ³⁰]
Rater 1 und Musterlösung	0.76	90.00%	0.93 [0.86, 1.00]
Rater 2 und Musterlösung	1.00	100.00%	1.00 [1.00, 1.00]
Rater 3 und Musterlösung	1.00	100.00%	1.00 [1.00, 1.00]
Rater 4 und Musterlösung	0.90	96.67%	0.92 [0.78, 1.00]
Durchschnitt Rater 1-4 und Musterlösung	0.92	96.67%	0.96 [0.91, 1.00]

Tabelle 38: Interraterreliabilität und konkurrente Validität, Items 24-29

Grundlage	Kappa	Prozentuale Übereinstimmung	AUC [95% KI]
Rater 1 und Musterlösung	0.53	76.67%	0.77 [0.64, 0.90]
Rater 2 und Musterlösung	0.73	86.67%	0.87 [0.75, 0.98]
Rater 3 und Musterlösung	0.73	86.67%	0.87 [0.75, 0.98]
Rater 4 und Musterlösung	0.40	70.00%	0.70 [0.54, 0.86]
Durchschnitt Rater 1-4 und Musterlösung	0.60	80.00%	0.80 [0.67, 0.93]

Insgesamt kann die VL „Therapie“ somit als geeignetes Instrument zur Erkennung von kritischen Therapieverläufen eingestuft und in ihrer jetzigen Form in der Vollzugspraxis eingesetzt werden.

²⁹ Ein Artefakt aufgrund der „Prävalenz“ der Antwortmöglichkeiten (d.h. Verteilung von Ja- und Nein-Antworten) kann ausgeschlossen werden, da die Verteilung bei der Auswertung von Items 24-29 ausgewogener ist, als bei derjenigen von Items 20-29. Da eine grössere Ausgewogenheit zu einem besseren Kappa führt und dies für Items 24-29 nicht vorliegt, können inhaltliche Gründe angenommen werden.

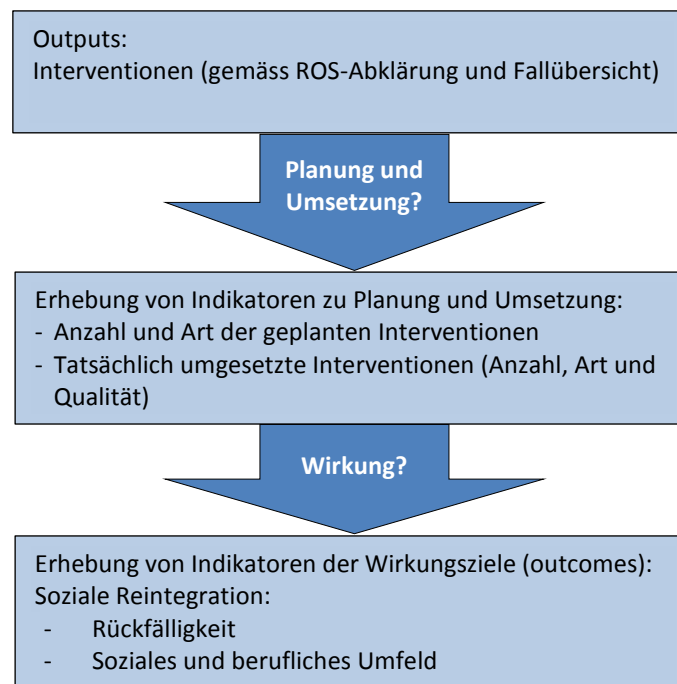
³⁰ KI = asymptotisches Konfidenzintervall. Das Konfidenzintervall gibt an, in welchem Intervall die AUC der Grundpopulation mit einer Wahrscheinlichkeit von 95% liegt.

7 Konzept für eine prospektive Studie

Mit ROS wurde erstmals der Versuch unternommen, den Sanktionenvollzug in einem strukturierten Prozessmodell zu regeln. Das Konzept der Risikoorientierung zieht sich dabei durch sämtliche Prozessstufen. Durch die retrospektive Studie konnte die erste und für den weiteren Prozess wegweisende Stufe, die Triagierung mittels FaST, evaluiert werden. Die Resultate aus Kapitel 5 deuten darauf hin, dass mit dem FaST Fälle ausfindig gemacht werden können, die einen erhöhten Abklärungsbedarf bezüglich Gewalt- und Sexualdelinquenz aufweisen.

In einem nächsten Schritt ist es zentral zu wissen, wie sich der gemäss Prozessmodell strukturierte Vollzug auf die soziale Reintegration von straffällig gewordenen Personen auswirkt („Wirksamkeit“ von ROS). Diese Überprüfung ist unabdingbar für eine vollständige Evaluation von ROS und sollte anhand von aktuellen, d.h. neu in das Vollzugssystem eintretenden Fällen erfolgen, deren künftige Entwicklung systematisch erfasst und evaluiert wird. Nachfolgend wird ein Konzeptvorschlag für eine solche *prospektive Studie* dargelegt.³¹ Allerdings handelt es sich dabei mehr um eine grobe Übersicht der relevanten Evaluationsbereiche denn um ein Detailkonzept. Anhand von Abbildung 7 werden die zentralen Fragestellungen erläutert.

Abbildung 7: Planung, Umsetzung und Wirkung von ROS-Interventionen



³¹ Für die konkrete Planung dieser Studie kann das „logische Modell“ (WKKF, 2001) herangezogen werden. Es stellt ein einfaches Modell zur Identifikation von Programmzielen und entsprechenden Erhebungsschritten (Indikatoren) dar. Darin wird zwischen Inputs/Ressourcen, Aktivitäten, Outputs und Outcomes unterschieden werden. Als Input können die Ressourcen des Programms verstanden werden, d.h. alles, was das Projekt ermöglicht, wie Geld, Arbeitszeit, Know-How etc. Als Output wird all das bezeichnet, was das Projekt zur Realisierung umsetzt. Hierzu kann Folgendes gezählt werden: ROS-Interventionen zur Steuerung des Vollzugs (FaST, diverse Verlaufslisten etc.) und sonstige bisherige klientel-orientierte Interventionen wie Lernprogramme, Therapien, Bewährungshilfe. Die Outcomes bezeichnen die gewünschten Wirkungen (den „Impact“ des Projekts), wie die Verhinderung bzw. Reduktion von Rückfälligkeit und die soziale Integration von Straftätern.

Die Wirkung von ROS auf die soziale Reintegration ist davon abhängig, welche Interventionen in der ROS-Abklärung und der Fallübersicht geplant und wie diese umgesetzt wurden. Es ist deshalb wichtig, nicht bloss die Wirkung, sondern auch die Art und Qualität von Planung und Umsetzung der ROS-Interventionen zu untersuchen (siehe die beiden dunkelblauen Pfeile in Abbildung 7). Für sämtliche Erhebungen ist es zentral, dass *messbare Indikatoren* erfasst werden. Die Evaluation sollte anhand einer Stichprobe aktueller Fälle durchgeführt werden.

In Kapitel 7.1 wird ein Vorschlag zur Ziehung der Stichprobe dargelegt, anhand derer die Komponenten (Planung/Umsetzung und Wirkung der Interventionen) untersucht werden können. In Kapitel 7.2 wird auf Aspekte zur Erhebung der Interventionsplanung (outputs) eingegangen, während sich Kapitel 7.3 mit der Erhebung der Wirkungen (outcomes) des ROS-Regimes beschäftigt.

7.1 Stichprobenziehung

Die Stichproben der retrospektiven und der prospektiven Studie sollten sich möglichst ähnlich sein, wenn man diese miteinander vergleichen möchte. Während die retrospektive Stichprobe auf abgeschlossenen Fällen beruhte, sollten für die prospektive Studie aktuelle, d.h. nicht abgeschlossene Fälle untersucht werden. Aktuelle Fälle haben den Vorteil, dass die zu erhebenden Daten direkt zugänglich sind, d.h. auf der Basis des Vollzugsverlaufs erhoben werden können, und man sich nicht auf teilweise lückenhafte Einträge in Akten oder Datenbanken stützen muss.

Idealerweise würde ein Untersuchungszeitraum von zwei Jahren (wie bei der retrospektiven Studie) definiert, in welchem laufend Angaben erfasst würden (z.B. von 1.1.2014 bis 31.12.2015). Da im Rahmen der retrospektiven Studie eine Vollerhebung für jene Personen mit einem G/S-Delikt gemäss FaST-Manual (siehe Anhang A, Seite 82) gemacht wurde, wäre für die prospektive Studie eine Vollerhebung aller C-Fälle, für welche während den zwei definierten Jahren in den BVD ein Geschäft eröffnet wurde, ebenfalls sinnvoll. Weil nur C-Fälle gemäss ROS-Prozessmodell behandelt werden, sind diese von besonderem Interesse. Aufgrund der ambivalenten Resultate der B-Fälle bei den Auswertungen der retrospektiven Studie (manchmal sind sie den A-Fällen näher, manchmal den C-Fällen), wäre es wünschenswert, zusätzlich genügend grosse Stichproben von A- und B-Fällen zu ziehen (hierzu wäre eine Poweranalyse zur Bestimmung der erforderlichen Stichprobengrösse durchzuführen). Dies müsste anhand einer Zufallsauswahl geschehen, welche sicherstellt, dass der gesamte Erhebungszeitraum von zwei Jahren und alle Abteilungen der BVD in ähnlichem Ausmass vertreten sind.

7.2 Planung und Umsetzung der ROS-Interventionen

Wie aus Abbildung 7 hervorgeht ist es wichtig, die geplanten ROS-Interventionen systematisch zu erfassen und zu deklarieren, ob und wie diese durchgeführt wurden.

Ein systematisches Vorgehen beinhaltet zunächst die Definition von *messbaren Indikatoren* für die Art und Qualität des individuellen Sanktionenvollzugs sowie später auch für die Messung des Erfolgs. Weiter ist es nötig genau zu dokumentieren, welche Interventionen ein Straftäter im Verlauf seiner Verbüßung der Sanktion absolviert. Hierbei ist einerseits die Anzahl der Interventionen, wie auch andererseits die Qualität der entsprechenden Intervention zu messen. Letzteres ist natürlich um einiges schwieriger zu erfassen, da dies subjektive Urteile z.B. der Therapeuten erfordert, die von Urteilen anderer Therapeuten stark abweichen können. Trotzdem sollten Anstrengungen unternommen werden, um herauszufinden, wie man die Qualität der Interventionsumsetzung messbar machen kann. Denn die Qualität der Programmanwendung („program delivery“) ist ein zentraler Einflussfaktor auf die ROS-Wirkungsziele. Aus der Evaluationsforschung ist dies hinlänglich bekannt: Das am besten konzipierte Programm ist nutzlos, wenn es schlecht angewandt bzw. umgesetzt wird.

Ein möglicher Weg, die Qualität der Programmanwendung zu überprüfen, wären *qualitative Interviews* von Tätern, deren Bewährungshelfern, aber auch Informanten aus dem sozialen Umfeld wie Angehörige oder Arbeitgeber. Sinnvoll wäre dies bei all jenen C-Fällen, für welche eine ROS-Abklärung erstellt wurde. Erfragt würden insbesondere Aspekte, welche sich nicht aus Akten oder Datenbanken entnehmen lassen wie beispielsweise das Ausmass an Motivation und Engagement bei der Arbeit, der Zuverlässigkeit im Rahmen der Zusammenarbeit mit der Bewährungshilfe oder das Nachgehen von Freizeitbeschäftigungen.

7.2.1 Planung der Interventionen

Bei C-Fällen geschieht die Planung von Interventionen anhand der ROS-Abklärung, welche dann durch den Fallverantwortlichen in der Fallübersicht ergänzt wird. Bei A- und B-Fällen existiert unseres Wissens kein vergleichbares Instrumentarium zur Planung von Interventionen. Vielmehr werden diese im RIS und in den Akten aufgeführt.

Zur systematischen Erfassung der Interventionen, unabhängig von der FaST-Kategorisierung, wäre es denkbar, dass der Fallverantwortliche eine Art Checkliste pro Fall zur Verfügung hat, auf welcher er die Interventionen, welche geplant wurden, ankreuzen kann. Angesichts der zahlreichen möglichen Interventionen, welche u.a. von der Tätergruppe, Deliktgruppe und der Institution, welche die Intervention durchführt, abhängt, müssten diese thematisch gebündelt werden. Die Checkliste würde idealerweise in elektronischer Form vorliegen, damit statistische Analysen zu Zusammenhängen mit anderen Variablen (z.B. FaST-Kategorisierung oder Rückfälligkeit) leichter durchgeführt werden könnten.

7.2.2 Umsetzung der Interventionen

Zur Überprüfung, inwiefern die geplanten Interventionen umgesetzt wurden, stellt das ROS-Prozessmodell mit den ROS-VL selbst ein Instrumentarium zur Verfügung. Die VL „Therapie“ (und die integrierte VL „Alarm“) ermöglicht es dem Fallverantwortlichen zu erkennen, inwiefern die in der ROS-Abklärung und Fallübersicht definierten Interventionsbereiche in einer ambulanten oder statio-

nären Therapie thematisiert wurden, und dient somit, neben der zentralen Funktion zur Erkennung von Hinweisen auf kritische Therapieverläufe, als Qualitätssicherungsinstrument bezüglich Interventionen, welche eine Therapie betreffen.

Interventionen, welche andere Bereiche als jenen der Therapie betreffen, werden ebenfalls durch ROS-VL abgedeckt. So existieren vier weitere VL, welche die Bereiche Vollzug, Bewährungshilfe, Interventionen und Motivation abdecken. Es ist zu prüfen, inwieweit messbare Indikatoren zur Umsetzung der geplanten Interventionen in verschiedenen Bereichen bereits vorliegen oder ggf. mit entsprechenden Indikatoren versehen werden sollten.

Dieser Qualitätssicherungszirkel, welcher in das ROS-Prozessmodell integriert ist, sollte eine längerfristige Überprüfung der Umsetzung der geplanten Interventionen ermöglichen.

An dieser Stelle sei nochmals darauf hingewiesen, dass die VL „Vollzug“, „Bewährungshilfe“, „Interventionen“ und „Motivation“ einer Evaluation im Sinne einer Reliabilitäts- und Validitätsprüfung bedürfen, wie diese für die VL „Therapie“ stattfand (siehe Kapitel 5.2).

7.3 Wirkung des ROS-Prozessmodells: Soziale Reintegration

Ein zentrales Ziel von ROS ist es, den Fokus der Vollzugsarbeit auf die *Rückfallprävention* zu setzen (Mayer & Treuthardt, 2011). Die Erfassung der Wirksamkeit der Vollzugsarbeit kann somit über eine Erhebung der Rückfälligkeitrate geschehen (Kapitel 7.3.1). Ob eine Person rückfällig wurde oder nicht steht in einem engen Zusammenhang mit der sozialen Reintegration: Wer sozial integriert ist, wird weniger rückfällig, so die Annahme. Das Ausmass der sozialen Reintegration kann jedoch auch anhand anderer Marker, insbesondere zum sozialen und beruflichen Umfeld, erhoben werden, was in Kapitel 7.3.2 thematisiert wird.

7.3.1 Rückfälligkeit

Im Rahmen der retrospektiven Studie wurde eine Basisrate der Rückfälligkeit erhoben, wobei Time-at-risks von einem, drei und sechs Jahren berücksichtigt wurden (Kapitel 5.2.6). Die Auswertungen erfolgten zudem differenziert nach A-, B- und C-Fällen. Die prospektive Studie sollte sich in Bezug auf die zu erhebenden Variablen an der retrospektiven Studie ausrichten. Konkret sollten zur Bestimmung der Rückfälligkeit die Variablen zum Anlassdelikt (Kapitel 5.1.4.3), zur Time-at-risk (Kapitel 5.1.4.4) und zu den Rückfällen (Kapitel 5.1.4.5) erhoben werden.

Beim Vergleich der prä-ROS-Fälle (retrospektiv) mit den ROS-Fällen (prospektiv) ist zu beachten, dass Veränderungen der Rückfälligkeit nicht alleine auf ROS zurückzuführen sein werden. Andere Einflussfaktoren, sog. Drittvariablen, wie beispielsweise die wirtschaftliche und politische Lage einer Gesellschaft (Lüdemann & Ohlemacher, 2002), die Anpassung des Strafgesetzbuches im Jahre 2007, allfällige Veränderungen in der Führung von Datenbanken wie dem VOSTRA, können einen Einfluss sowohl auf den Input (Art der Urteile) des Vollzugssystems, als auch auf die registrierten Rückfälle haben. Dies ist für die künftige Einschätzung der Wirksamkeit von ROS, soweit möglich und plausibel, unbedingt zu berücksichtigen.

Laut Analysen des Bundesamtes für Statistik (2011) muss bei der Erhebung von Rückfälligkeitsraten unbedingt beachtet werden, dass diese insbesondere im ersten Jahr nach einer Straftat stark unter-

schätzt werden. Dies hat damit zu tun, dass Verurteilungen zwar im Strafregister erfasst werden, sobald ein Urteil in Kraft tritt, die Abhandlung von allfälligen Rechtsmittelverfahren jedoch mehrere Monate oder gar Jahre dauern kann. Die Strafregistereinträge eines Jahres werden somit mit der Zeit genauer. Dementsprechend ist es wichtig, möglichst lange mit der Erhebung der Rückfälligkeit der prospektiven Fälle zu warten bzw. nach einigen Jahren eine Kontrolle und entsprechende Anpassungen vorzunehmen.

Würde ein Untersuchungszeitraum von Fällen der Jahre 2014 und 2015 vorgenommen, müsste die Rückfälligkeit nach einem Jahr Time-at-risk (2015 bzw. 2016) idealerweise bei Abschluss der prospektiven Studie (im Jahre 2021, wenn ein Untersuchungszeitraum von sechs Jahren gewählt würde), kontrolliert und angepasst werden.

Aufgrund der oben beschriebenen Drittvariablen und somit entstehenden Einschränkung bezüglich der Vergleichbarkeit von prä-ROS- und ROS-Rückfallraten bestünde eine Möglichkeit für eine verlässlichere Einschätzung der Wirksamkeit darin, dass man einen Vorher-Nachher-Vergleich zwischen dem ROS-Kanton Zürich und einem Nicht-ROS-Kanton durchführen würde. Drittvariablen könnten dadurch besser kontrolliert werden (Schnell, Hill & Esser, 2005). Ein solches Design hätte jedoch den Rahmen der Evaluation des Modellversuchs gesprengt, da die Datenerhebung einen hohen zeitlichen Aufwand mit sich bringt. Ebenfalls wurde die Machbarkeit zur Durchführung eines kontrollierten Experiments (im Sinne von zufälliger Zuweisung von Gewalt- und Sexualstraftätern auf die Bedingungen „ROS-Abklärung“ und „keine ROS-Abklärung“) seitens des Auftraggebers als nicht gegeben erachtet.

7.3.2 Soziales und berufliches Umfeld

Nicht nur die Rückfälligkeit, sondern auch das soziale und berufliche Umfeld fungiert als Indikator für die soziale Reintegration. Das steuerbare Einkommen wurde bereits in der retrospektiven Studie erhoben. Das Vorhandensein eines stabilen Einkommens deutet darauf hin, dass eine Person einer Erwerbstätigkeit nachgeht, womit ein wichtiger Faktor der sozialen Integration gegeben ist. Zusätzlich zum steuerbaren Einkommen wäre es wünschenswert, den Beziehungsstatus einer Person zu erfassen. Die Forschung zeigt, dass eine stabile Partnerschaft ein protektiver Faktor für Delinquenz darstellt (z.B. Andrews & Bonta, 2010; Sampson et al., 2006). Einen groben Indikator für den Beziehungsstatus einer Person liefert der Zivilstand, welcher im Rahmen der prospektiven Studie ebenfalls erfasst werden könnte. Das steuerbare Einkommen und der Zivilstand sollten parallel zur Erhebung der Rückfälligkeit erfasst werden (idealerweise in Anlehnung an die retrospektive Studie ein Untersuchungszeitraum von sechs Jahren nach Beginn der Time-at-risk).

Wie sich bei der Erhebung der kantonalen Steuerdaten für die retrospektive Studie gezeigt hat, ist es äusserst schwierig, die Personen im Nachhinein ausfindig zu machen. Ein Vergleich der prä-ROS-Fälle mit den ROS-Fällen hinsichtlich des steuerbaren Einkommens wird nur unter dem Vorbehalt der vielen Ausschlüsse in der Stichprobe der retrospektiven Studie möglich sein. Sofern diese Quelle berücksichtigt werden soll, muss eine bessere Datenlage durch detailliertere Erhebungen ermöglicht werden. Interessant wäre im Rahmen der prospektiven Studie jedoch eine Analyse zum Zusammenhang zwischen Steuerbarem Einkommen, Zivilstand und Rückfälligkeit.

8 Zusammenfassung und Schlussfolgerungen

Den Sanktionenvollzug im Rahmen eines klar strukturierten Prozessmodells zu steuern – das ist das Hauptanliegen des Modellversuchs ROS.

Die vorliegende Ergebnisevaluation verfolgte dabei folgende Ziele:

1. Die Prüfung der Güte des Triage-Instruments FaST
2. Die Erhebung einer Basisrate der Rückfälligkeit von Straftätern, deren Sanktion vor Einführung von ROS durch das Amt für Justizvollzug des Kantons Zürich administriert wurde
3. Die Prüfung der Güte der Verlaufsliste Therapie/Alarm

Nachfolgend werden die zentralen Ergebnisse und Erkenntnisse aus der Ergebnisevaluation zusammengetragen und Schlussfolgerungen für den Einsatz von ROS in der Vollzugspraxis gezogen.

8.1 Überprüfung der Güte des Triage-Instruments FaST

Die Überprüfung der Güte des Triage-Instruments FaST umfasste die Überprüfung der Reliabilität und der Validität. Die Reliabilität, das heisst der Grad der Genauigkeit des FaST, wurde anhand der **Interraterreliabilität** erfasst.

- *Im Rahmen der Interraterreliabilitätsstudie konnte gezeigt werden, dass das FaST über drei Rater sehr gute Übereinstimmungen vorbrachte.*

Dies deutet darauf hin, dass das FaST bei seiner Durchführung wenig Interpretationsspielraum zulässt. Gerade weil das FaST zu Beginn des Vollzugs eingesetzt wird und über den weiteren Abklärungsbedarf einer Person bestimmt, ist es wichtig, dass es unabhängig vom Anwender zum gleichen Resultat führt.

Unter Validitätsaspekten wurde im Rahmen der retrospektiven Studie untersucht, ob das FaST tatsächlich jene Personen identifiziert, die einen erhöhten Abklärungsbedarf bezüglich G/S-Delinquenz aufweisen. Zur Überprüfung der **prädiktiven Validität** wurde insbesondere der Zusammenhang zwischen dem FaST und der Rückfälligkeit untersucht. Dabei interessierten vorwiegend C-Fälle, da diese auf einen erhöhten Abklärungsbedarf bezüglich Gewalt- und Sexualdelinquenz hindeuten. Neben der Rückfälligkeit wurde als weiteres Kriterium das steuerbare Einkommen nach verschiedenen Time-at-risks herangezogen. Sowohl die Rückfälligkeit, als auch das steuerbare Einkommen fungieren als Indikatoren für die soziale Integration einer Person. Folgende Ergebnisse und Erkenntnisse zur prädiktiven Validität sind zentral:

- C-Fälle werden *signifikant häufiger* mit einem G/S-Delikt rückfällig als A- und B-Fälle.
- Das FaST erkennt vornehmlich jene Personen als abklärungsbedürftig, die bereits einmal ein G/S-Delikt begangen haben, verkennt jedoch jene Fälle, die noch nie mit einem G/S-Delikt aufgefallen sind. Hierbei sind zwei Schlüsse möglich: (1) Man anerkennt diese Schwäche, die angesichts des Umstands, dass Täter, die zum ersten Mal mit einem G-Delikt straffällig wurden, dies mit einem leichteren Gewaltdelikt tun. (2) Wollte man das FaST verbessern, so dass es besser in der Lage ist Personen zu erkennen, die noch nie ein G/S-Delikt begangen haben, jedoch ein erhöhtes Risiko für die Begehung eines solchen aufweisen, müsste man andere Faktoren neu hinzufügen oder stärker gewichten, die sich nicht auf das Anlassdelikt bzw. die Vorstrafen beziehen. Denkbar wäre, dass sich das FaST, ähnlich dem LSI-R (Level of Service Inventory Revised; Andrews & Bonta, 2010), vermehrt auf andere Faktoren wie Einstellungen, soziales Umfeld etc. stützen würde. Dadurch würde sich das FaST zu einem ausführlicheren Instrument wandeln und der praktische Nutzen eines schnellen Screening-Instruments ginge verloren.
- Einschränkung: Unsere Stichprobe enthielt zu wenig Fälle mit einem Sexualdelikt als Anlassdelikt für detailliertere Analysen.
- Das FaST scheint zwischen Gewaltdelikten mit tiefem und hohem Strafmass differenzieren zu können: Personen, die mit einem G1-Delikt (tiefes Strafmass) rückfällig wurden, wurden vom FaST meist als A-Fall gewertet. Bezüglich Personen, die mit einem G3-Delikt (höchstes Strafmass) rückfällig wurden, bestand die Tendenz, dass das FaST sie als C-Fall klassifizierte (vier von fünf Fällen). Angesichts der sehr wenigen G3-Rückfälle ist die Aussagekraft jedoch sehr beschränkt.
- C-Fälle werden **nicht** schneller mit einem G/S-Delikt rückfällig als A- und B-Fälle. Wenn A/B-Fälle mit einem Gewaltdelikt rückfällig werden, handelt es sich meistens um ein leichteres Gewaltdelikt. Ersteres könnte mit einer engmaschigeren Begleitung durch die Bewährungshilfe erklärt werden oder auch mit typischen Merkmalen der jeweiligen Populationen (z.B. Personen, die durch Beschaffungskriminalität ihre Drogensucht finanzieren).
- Mehrere Resultate zeigen, dass sich B-Fälle hinsichtlich verschiedener Aspekte von den A- und C-Fällen unterscheiden. Dies deutet darauf hin, dass die B-Fälle eine distinkte Kategorie bilden, wie dies ebenfalls in der Konzeption des FaST vorgesehen ist (B-Fälle sollten demnach einen besonderen Abklärungsbedarf bezüglich allgemeiner Delinquenz aufweisen). Folgende Resultate zu den B-Fällen sind zentral:
 - Mit einem G/S-Delikt rückfällig gewordene B-Fälle waren häufig Personen, die bereits einmal ein G/S-Delikt begangen haben (Effekt knapp nicht signifikant). A-Fälle hingegen, die mit einem G/S-Delikt rückfällig wurden, haben ein solches Delikt meist zum ersten Mal begangen (signifikanter Effekt).
 - B-Fälle weisen die längste mittlere Time-at-risk bis zum ersten Rückfall auf, wenn man alle Rückfallarten betrachtet. Der Unterschied ist für die B- versus C-Fälle signifikant.
 - Nach einer time-at-risk von 6 Jahren werden B-Fälle nicht häufiger mit einem allgemeinen Delikt rückfällig, wie dies von der FaST-Konzeption vorgesehen wäre. Dies könnte, wie bereits ausgeführt, mit einer fehlenden Differenzierung der „allgemeinen Delikte“ zu tun haben.



- *B-Fälle werden (knapp nicht signifikant) häufiger erst nach regulärem Haftende entlassen als A- und C-Fälle. Dies deutet darauf hin, dass B-Fälle Problembereiche aufweisen, die eine vorzeitige bedingte Entlassung verhindern.*
- *Die Auswertungen zum steuerbaren Einkommen als Indikator für die soziale Reintegration zeigen, dass B-Fälle über alle untersuchten Time-at-risks das kleinste Einkommen aufweisen, was wiederum für eine spezifische Problematik dieser Personen spricht.*

Daraus wird deutlich, dass ein Überdenken der gängigen ROS-Praxis, nämlich A- und B-Fälle gleich zu behandeln, notwendig ist und an einer ressourcenorientierten Lösung gearbeitet werden sollte. Um adäquate Interventionsformen für die B-Fälle zu finden, wäre es zudem sinnvoll, die allgemeinen Delikte weiter zu differenzieren. Insbesondere wäre es nötig zu prüfen, inwiefern Verstösse gegen das Ausländergesetz (AuG) für eine B-Kategorisierung verantwortlich sind.

- *Weder die Höhe des steuerbaren Einkommens, noch das blosse Vorhandensein eines Einkommens zeigten konsistente Zusammenhänge mit der FaST-Kategorisierung auf. Dies hängt vermutlich mit dem Einfluss von nicht-erhobenen Drittgrössen zusammen, wie z.B. Verbüsung von Freiheitsstrafen aufgrund von Rückfällen oder sozialen Umständen wie einer stabilisierenden Partnerschaft.*

Zusätzlich zur Überprüfung der prädiktiven Validität wurde für die Evaluation die **konvergente Validität** herangezogen. Da der Fokus der Evaluation auf den C-Fällen liegt, wurde ein Instrumentarium zur Einschätzung des Risikos für Gewalt- und Sexualstraftaten herangezogen. Als solches bot sich der VRAG. Folgende Ergebnisse und Erkenntnisse sind zentral:

- *Der VRAG konnte lediglich bei 13% aller Fälle angewendet werden (häufig kein VRAG-relevantes G/S-Delikt oder fehlende Informationsgrundlage zur Durchführung eines VRAG).*
- *Der VRAG konnte für eine beachtliche Anzahl A- und B-Fälle ausgefüllt werden.*
- *A-Fälle wiesen gegenüber B- und C-Fällen einen signifikant tieferen Mittelwert der VRAG-Risikoklassen auf.*
- *Bei A- und B-Fällen, die einen VRAG zuliessen, handelte es sich häufig um das Anlassdelikt der einfachen Körperverletzung, welche im Gegensatz zu den anderen G/S-Delikten als relativ mild einzustufen ist.*
- *C-Fälle stellen eine Gruppe von Verurteilten dar, die in Bezug auf das Rückfallrisiko gemäss VRAG sehr heterogen ist: Die Streuung der VRAG-Risikoklassen ist innerhalb der C-Fälle breit.*

Das Resultat, dass lediglich bei 13% aller Fälle ein VRAG angewendet werden konnte, macht deutlich, dass die konvergente Validität idealerweise anhand eines Instrumentariums erhoben werden sollte, welches dem FaST ähnlicher ist. Konkret sollte es sich um ein Instrument handeln, welches für eine ähnlich grosse Gruppe von G/S-Tätern anwendbar ist wie das FaST. So ist der VRAG lediglich bei einer kleinen Gruppe von G/S-Tätern anwendbar. Das FaST hingegen weist einen viel längeren Katalog an G/S-Delikten auf, die speziell gewichtet werden.

Da jedoch unseres Wissens kein Instrument existiert, welches dem FaST auf inhaltlicher und formeller Ebene ähnlich ist, schien der VRAG zur Überprüfung der konvergenten Validität hinreichend geeignet, zumal der Fokus von ROS auf die G/S-Straftäter gerichtet ist.

Für differenziertere Aussagen zur konvergenten Validität von FaST und VRAG wären jedoch Analysen auf der Grundlage einer repräsentativeren Stichprobe nötig. Eine solche könnte besser anhand von aktuellen Fällen vorgenommen werden. Demnach würde es sich anbieten, die konvergente Validität von FaST und VRAG erneut im Rahmen der prospektiven Studie zu untersuchen.

8.2 Erhebung einer Basisrate der Rückfälligkeit von prä-ROS-Fällen

Für einen Vergleich der Rückfälligkeit von prä-ROS- und ROS-Fällen wurde im Rahmen der retrospektiven Studie die Rückfälligkeit von prä-ROS-Fällen nach einem, drei und sechs Jahren Time-at-risk untersucht, wobei folgende Ergebnisse zentral sind:

- Nach einem Jahr Time-at-risk wurden bereits 34.7% aller Personen rückfällig, bei den C-Fällen sind es gar 54.7%.
- Die Rückfälligkeit mit den Ausprägungen „kein Rückfall“, „Rückfall allgemeine Delinquenz“ und „Rückfall G/S-Delinquenz“ weist nach einem Jahr und sechs Jahren Time-at-risk einen signifikanten Zusammenhang mit dem FaST auf. Nach drei Jahren ist der Effekt knapp nicht signifikant.
- Insbesondere werden C-Fälle nach einer Time-at-risk von sechs Jahren häufiger mit einem G/S-Delikt rückfällig als A- und B-Fälle.

Könnte in der prospektiven Studie gezeigt werden, dass die Rückfälligkeit von ROS-C-Fällen geringer ausfällt als jene von prä-ROS-C-Fällen, würde dies für die Wirksamkeit des ROS-Prozessmodells sprechen.

8.3 Überprüfung der Güte der ROS-Verlaufsliste „Therapie“

Die Güte der ROS-VL „Therapie“ wurde anhand der Interraterreliabilität und der konkurrenten Validität untersucht. Die **Interraterreliabilitätsprüfung** brachte folgende Resultate hervor:

- *Von 29 Items wiesen 24 eine gute, 1 eine befriedigende und 4 eine unbefriedigende Übereinstimmung auf.*
- *Die 4 Items mit unbefriedigenden Interraterreliabilitäten wurden zusammen mit dem ROS-Entwicklungsteam diskutiert, wobei die Gründe für die unbefriedigenden Resultate mit hinreichender Wahrscheinlichkeit ausfindig gemacht werden konnten. Anpassungen der VL zur Behebung dieser Unklarheiten wurden vorgenommen.*
- *Es konnte zudem gezeigt werden, dass Fälle mit Hinweisen auf einen kritischen Therapieverlauf übereinstimmend erkannt wurden.*

Zur Überprüfung der **konkurrenten Validität** wurde überprüft, inwiefern die vier Rater Hinweise auf kritische Therapieverläufe übereinstimmend mit Musterlösungen („wahre Werte“) als solche erkannten. Es sind folgende Resultate zu konstatieren:

- *Die konkurrente Validität, die sich auf die Erkennung eines kritischen Therapieverlaufs bezieht (KEIN ZWINGENDER Eintrag in die Alarmliste), ist mit einer AUC = 0.96 als sehr gut zu bewerten.*
- *Die konkurrente Validität, die sich auf die Erkennung eines kritischen Therapieverlaufs bezieht (ZWINGEND Eintrag in die Alarmliste), ist mit einer AUC von 0.80 als gut einzustufen.*
- *Insgesamt kann die VL „Therapie“ als geeignetes Instrument zur Erkennung von kritischen Therapieverläufen eingestuft und in ihrer jetzigen Form in der Vollzugspraxis eingesetzt werden.*

8.4 Ausblick auf prospektive Studie

Mit der retrospektiven Studie konnte die Güte des Triage-Instrumentes FaST überprüft werden. Um eine Aussage zur *Wirksamkeit* von ROS machen zu können, ist längerfristige Analyse von ROS-Fällen nötig. Diese kann im Rahmen einer prospektiven Studie durchgeführt werden, bei der eine vergleichbare Stichprobe von Fällen untersucht wird.

Ziel ist es festzustellen, wie sich ROS auf die soziale Reintegration von straffälligen Personen auswirkt. Als messbare Indikatoren für die soziale Reintegration dienen sowohl *Rückfallzahlen*, als auch andere soziale Indikatoren, wie etwa *Angaben zu Erwerbstätigkeit, Beziehungsstatus einer Person (bzw. Fähigkeit soziale Beziehungen einzugehen)*. Ein Vergleich mit den retrospektiven Fällen lässt eine gewisse Einschätzung der Wirksamkeit von ROS zu. Dabei ist unbedingt zu beachten, dass auch andere Variablen einen nicht messbaren Einfluss auf die Rückfälligkeit haben können (wirtschaftliche

und politische Lage, Anpassung des Strafgesetzbuches im Jahre 2007 oder Veränderungen in der Führung von Datenbanken wie dem VOSTRA). Nach allgemein anerkannten wissenschaftlichen Standards ist für eine verlässliche Einschätzung der Wirksamkeit von Interventionen jedoch nicht nur ein Vorher-Nachher-Vergleich, sondern auch der Vergleich mit der Entwicklung einer geeigneten „Kontrollgruppe“ nötig. Die Wirksamkeit könnte also verlässlicher eingeschätzt werden, wenn die Entwicklung im Kanton Zürich mit einem oder mehreren anderen Nicht-ROS-Kantonen verglichen werden könnte. Dazu wären aber Daten zur Rückfälligkeit in diesem Kanton nötig, die einen ähnlichen Vorher-Nachher-Vergleich zulassen.

Die Wirksamkeit ist in hohem Mass davon abhängig, wie die ROS-Interventionen *geplant* und *umgesetzt* werden. Auch hier ist darauf zu achten, messbare Indikatoren zu definieren. Neben der Erhebung von Anzahl und Art der geplanten Interventionen, muss in einem zweiten Schritt eruiert werden, ob die geplanten Interventionen auch *umgesetzt* wurden. In einem dritten Schritt ist es unabdingbar, die *Qualität* der umgesetzten Interventionen zu eruieren. Denkbar wären hier qualitative Interviews von Tätern, deren Bewährungshelfern und Personen aus dem Umfeld des Täters.

Wünschenswert wäre ein permanenter Qualitätssicherungszyklus, welcher die systematische Erfassung genau definierter, messbarer Indikatoren festhält. Die Erfassung dieser Indikatoren sollte idealerweise direkt im ROS-Prozessmodell verankert sein.

8.5 Fazit

Unsere Ergebnisevaluation zeigte, dass das **FaST**, als eines der zentralen ROS-Instrumente, ein vielversprechendes Instrument darstellt. Es ist gemäss der vorliegenden Evaluation zuverlässig und valide. Die Aussagekraft der vorgenommenen Validitätsprüfung ist jedoch durch die stark reduzierte Anwendbarkeit des VRAG als Vergleichsinstrument eingeschränkt. Dies kann in einer prospektiven Vergleichsanordnung mit einer systematischen Erhebung der VRAG-Variablen nachgeholt werden. Verbesserungspotential besteht ausserdem in Bezug auf eine bessere Abgrenzung der B-Fälle von den A- und C-Fällen.

Auch die **Verlaufslisten** Therapie bzw. Alarm erwiesen sich für Praxiszwecke als verlässlich und valide. Grundsätzlich ist es der Qualität von angeordneten Therapien sicherlich nur förderlich, wenn überprüft wird, ob die Arbeit an den Problembereichen eines Täters einen kritischen Verlauf annimmt, damit gegebenenfalls seitens der Vollzugsbehörde reagiert werden kann.

Somit sind wichtige Voraussetzungen nicht nur für den praktischen Einsatz von zwei wichtigen ROS-Instrumenten gegeben, sondern auch für die Möglichkeit, überhaupt Wirkungen im Sinne von wirksamen Therapien und mittel- bis langfristig reduzierter Rückfälligkeit zu erzielen.

9 Literatur

- Amt für Justizvollzug Kanton Zürich (2011). *Modellversuch Risikoorientierter Sanktionenvollzug (ROS) der Kantone Luzern, Thurgau, St. Gallen und Zürich – Zwischenbericht I: Entwicklungsstand im September 2011*. Zürich (unveröffentlichter Bericht).
- Amt für Justizvollzug Kanton Zürich (2012). *Modellversuch Risikoorientierter Sanktionenvollzug (ROS) der Kantone Luzern, Thurgau, St. Gallen und Zürich – Zwischenbericht II: Entwicklungsstand im September 2012*. Zürich (unveröffentlichter Bericht).
- Amt für Justizvollzug Kanton Zürich, Psychiatrisch-Psychologischer Dienst (2013). *Zusammenfassung der Ergebnisse zur Untersuchung der Inter-Rater-Reliabilität der „Verlaufsliste Therapie“ (ROS)*. Zürich (unveröffentlichtes Dokument).
- Andrews, D. A. & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). New Providence: LexisNexis.
- Baur, M., Manzoni, P. & Schwarzenegger, C. (2013). *Modellversuch Risikoorientierter Sanktionenvollzug (ROS) – Ergebnisevaluation: Zwischenbericht II*. Zürich (unveröffentlichter Bericht).
- Bewährungs- und Vollzugsdienste Zürich, Fachsupport und Lernprogramme (2010). *Risikoorientierter Sanktionenvollzug, Fall-Screening-Tool (FaST)*. Zürich (unveröffentlichtes Dokument).
- Bewährungs- und Vollzugsdienste Zürich, Fachsupport und Lernprogramme (2013). *Fall-Screening-Tool (FaST) Manual* (Version 3.0, Juni 2013). Zürich (unveröffentlichtes Dokument).
- Bewährungs- und Vollzugsdienste Zürich, Projekt- und Qualitätsmanagement (2012). *Projekt ROS: Information Betrieb ab 01. Juli 2012* (Stand 29.5.2012). Zürich (unveröffentlichtes Dokument).
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Bundesamt für Statistik (2011). *Neues Sanktionenrecht und strafrechtlicher Rückfall, erste Analysen der Rückfallentwicklung seit Inkraftsetzung des Allgemeinen Teils des Strafgesetzbuches im Jahr 2007*. In [EDJ](#), besucht am 23.7.2013.
- Endrass, J., Rossegger, A. & Urbaniok, F. (2007). *Zürcher Forensik Studie, Abschluss des Modellversuchs: Therapieevaluation und Prädiktorenforschung*. In [EJPD](#), besucht am 22.07.2013.
- Field, A. (2013). *Discovering statistics using SPSS*. Sage: Los Angeles.
- Gerrig, R. J. & Zimbardo, Ph. G. (2008). *Psychologie*. München: Pearson Studium.
- Gwet, L. G. (2010). *Handbook of inter-rater reliability*. Gaithersburg: Advanced Analytics.
- Hare, R. D. (1970). *Psychopathy: theory and research*. New York: John Wiley & Sons.
- Harris, G. T., Rice, M. E. & Quinsey, V. L. (1994). Psychopathy as a taxon: evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology*, 62(2), 387-397.

- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77-89.
- Hosmer, D. W. & Lemeshow, S. (2000). *Applied logistic regression*. New York: John Wiley & Sons.
- Hussy, W., Schreier, M. & Echterhoff, G. (2013). *Forschungsmethoden in Psychologie und Sozialwissenschaften*. Berlin: Springer.
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills: Sage.
- Lüdemann, Ch. & Ohlemacher, T. (2002). *Soziologie der Kriminalität, theoretische und empirische Perspektiven*. Weinheim: Beltz Juventa.
- Mayer, K. & Treuthardt, D. (2011). *Risikoorientierter Sanktionenvollzug (ROS), Grundlagen, Strategien und Instrumente* (nicht veröffentlichte Masterarbeit). Universität Zürich und Institut für Opferchutz und Täterbehandlung, Zürich.
- Mossmann, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62, 783-792.
- Powers, D. M. W. (2012). *The Problem with Kappa*. Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, 345–355.
- Quinsey, V. L., Harris, G. T., Rice, M. E. & Cormier, C. A. (2006). *Violent offenders: appraising and managing risk*. Washington: American Psychological Association.
- Rossegger, A., Urbaniok, F., Danielsson, C. & Endrass, J. (2009). Der Violence Risk Appraisal Guide (VRAG) - ein Instrument zur Kriminalprognose bei Gewaltstraftätern: Übersichtsarbeit und autorisierte deutsche Übersetzung. *Fortschritte der Neurologie und Psychiatrie* 77(10), 577-584.
- Sampson, R. S., Laub, J. H. & Wimer, C. (2006). Does marriage reduce crime? A counterfactual approach to within-individual causal effects. *Criminology*, 44(3), 465-508.
- Schnell, R., Hill, P. B. & Esser, E. (2005). *Methoden der empirischen Sozialforschung*. München: Oldenbourg Wissenschaftsverlag.
- Schwarzenegger, C., Manzoni, P. & Baur, M. (2012). *Modellversuch Risikoorientierter Sanktionenvollzug (ROS) – Ergebnisevaluation: Zwischenbericht I*. Zürich (unveröffentlichter Bericht).
- Sim, J. & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy*, 85(3), 257-268.
- Swets, J. A. (1988). Measuring the Accuracy of Diagnostic Systems. *Science*, 240(4857), 1285-1293.
- Urbaniok, F. (2007). *FOTRES – Forensisches Operationalisiertes Therapie- und Risiko-Evaluations-System*. Bern: Zytglogge.
- WKKF (W. K. Kellogg Foundation) (2004). *Logic Model Development Guide. Using Logic Models to Bring Together Planning, Evaluation, and Action*. In <http://www.epa.gov/evaluate/pdf/eval-guides/logic-model-development-guide.pdf>, besucht am 8.8.2013.

10 Anhang A: Gewalt- und Sexualdelikte sowie schwerwiegende SVG-Delikte nach FaST

Tabelle 39: Gewalt-, Sexualdelikte sowie schwere SVG-Delikte nach FaST-Manual

Gesetzesartikel	Gewichtung
Gewalt- und Sexualdelikte	
StGB2 111 Vorsätzliche Tötung	G3
StGB 112 Mord	G3
StGB 113 Totschlag	G2
StGB 114 Tötung auf Verlagen	G1
StGB 115 Beihilfe zum Selbstmord	G2
StGB 116 Kindstötung	G1
StGB 117 Fahrlässige Tötung	G1
StGB 118 Abtreibung durch die Schwangere	G2
StGB 122 Schwere Körperverletzung	G3
StGB 123 Einfache Körperverletzung	G1
StGB 125 Fahrlässige Körperverletzung	G1
StGB 125.2 Fahrlässige KV (schwere Schädigung)	G2
StGB 126.2 und folgende Tötlichkeiten	G1
StGB 127 Aussetzung	G2
StGB 128 Unterlassen der Nothilfe	G1
StGB 129 Gefährdung des Lebens	G2
StGB 133 Raufhandel	G1
StGB 134 Angriff	G2
StGB 135.1 Gewaltdarstellungen	G1
StGB 136 Verabreichen gesundheitsgefährdender Stoffe an Kinder	G1
StGB 140.1 Raub	G1
StGB 140.2 u. 3 qualifizierter Raub	G2
StGB 140.4 qualifizierter Raub	G3
StGB 156.1 bis 3 Erpressung	G2
StGB 156.4 Erpressung	G3
StGB 180 Drohung	G1
StGB 181 Nötigung	G1
StGB 183 Freiheitsberaubung und Entführung	G2
StGB 184 Freiheitsberaubung (erschwerende Umstände)	G3
StGB 185 Geiselnahme	G2
StGB 185.3 Geiselnahme (besonders schwerer Fall)	G3
StGB 187 Sexuelle Handlungen mit einem Kind	S2
StGB 187.4 Sex. Handl. mit einem Kind (irrige Vorstellung)	S1
StGB 188.1 Sex. Handl. mit Abhängigen	S1
StGB 189 Sexuelle Nötigung	S1
StGB 190 Vergewaltigung	S3
StGB 191 Schändung	S3



Gesetzesartikel	Gewichtung
StGB 192 Sex. Handl. mit Anstaltspfleglingen etc.	S1
StGB 193 Ausnützung einer Notlage	S1
StGB 194 Exhibitionismus	S1
StGB 196.1 und 3 Menschenhandel	S1
StGB 196.2 qualifizierter Menschenhandel	S2
StGB 197 Pornografie	S1
StGB 198 Sexuelle Belästigung	S1
StGB 213 Inzest	S2
StGB 221.1 Brandstiftung	G2
StGB 221.2 Brandstiftung	G3
StGB 221.3 Brandstiftung	G1
StGB 223.1.1 Verursachen einer Explosion	G2
StGB 223.1.2 Verursachen einer Explosion	G1
StGB 224 Gefährdung durch Sprengstoffe etc.	G2
StGB 225.1 Gefährdung durch Sprengstoffe etc.	G1
StGB 226.1 Herstellen etc. von Sprengstoffen usw.	G2
StGB 226.2 u. 3 Herstellen etc. von Sprengstoffen usw.	G1
StGB 227.1.1 u. .1.2 Verursachen einer Überschwemmung etc.	G1
StGB 228.1 Beschädigung von elektrischen Anlagen etc.	G1
StGB 231.1.1 Verbreiten menschlicher Krankheiten	G2
StGB 231.1.2 Verbreiten menschlicher Krankheiten	G1
StGB 234.1 Verunreinigung von Trinkwasser	G2
StGB 237.1 Störung des öffentlichen Verkehrs	G1
StGB 238.1 Störung des Eisenbahnverkehrs	G1
StGB 260BIS111, 112, 122 Vorbereitungshandlungen	G2
StGB 260BIS140, 183, 185, 221 Vorbereitungshandlungen	G1
StGB 260BIS264 Vorbereitungshandlungen	G3
StGB 260QUATER Gefährdung der öffentlichen Sicherheit mit Waffen	G1
StGB 260QUINQUIES Finanzierung des Terrorismus	G2
StGB 260TER Kriminelle Organisation	G2
StGB 261BIS Rassendiskriminierung	G1
StGB 262 Störung des Totenfriedens	G1
StGB 264 Völkermord	G3
StGB 285 Gewalt und Drohung geg. Behörden etc.	G1
StGB 311.1 Meuterei von Gefangenen	G1
StGB 311.2 Meuterei von Gefangenen	G2
A StGB3 139 Raub	G1
A StGB 139.1BIS Raub (Mitführen einer Waffe)	G2
A StGB 139.2 Qualifizierter Raub	G2
A StGB 156 Erpressung	G1
A StGB 182 Freiheitsberaubung	G1
A StGB 187 Notzucht	S2
A StGB 188 Nötigung zu einer andern unzüchtigen Handlung	S1
A StGB 189 Schändung	S2
A StGB 190 Unzucht mit Schwachsinnigen	S2



Gesetzesartikel	Gewichtung
A StGB 191.1 u. 2 Unzucht mit Kindern	S2
A StGB 191.3 Unzucht mit Kindern (irrig Vorstellung)	S1
A StGB 192 Unzucht mit unmündigen Pflegebefohlenen	S2
A StGB 193 Unzucht mit Anstaltspfleglingen etc.	S1
A StGB 194 Widernatürliche Unzucht	S2
A StGB 196.1 Menschenhandel	G2
A StGB 196.2 Menschenhandel (Anstalten treffen)	G1
A StGB 200 Begünstigung der Unzucht	G1
A StGB 201 Zuhälterei	G1
A StGB 202 Frauen- und Kinderhandel	G2
A StGB 203 Öffentliche unzüchtige Handlungen	G1
A StGB 204 Unzüchtige Veröffentlichungen	G1
A StGB 205 Unzüchtige Belästigung	S1
A StGB 206 Anlocken zur Unzucht	S1
A StGB 260BIS139 Vorbereitungshandlungen (zu Raub)	G1
A StGB 264 Tierquälerei	G1
AMSTG4 130 Raub	G1
AMSTG 153 Notzucht	S2
AMSTG 154 Nötigung zu einer anderen unzüchtigen Handlung	S1
AMSTG 157 Widernatürliche Unzucht	S2
AMSTG 159 Öffentliche unzüchtige Handlungen	S1
MSTG5 115 Vorsätzliche Tötung	G3
MSTG 116 Mord	G3
MSTG 120 Fahrlässige Tötung	G1
MSTG 121 Schwere Körperverletzung	G3
MSTG 122 Einfache Körperverletzung	G1
MSTG 128 Raufhandel	G1
MSTG 128A1 Angriff	G2
MSTG 149 Drohung	G1
MSTG 150 Nötigung	G1
MSTG 151 A1 und A2 Freiheitsberaubung	G1
MSTG 153 Sexuelle Nötigung	S1
MSTG 154 Vergewaltigung	S3
MSTG 155 Schändung	S3
MSTG 156.1 Sexuelle Handlungen mit Kindern	S2
MSTG 159 A Sexuelle Belästigungen	S1
MSTG 160 A1 und A3 Fahrlässige Verursachung einer Feuersbrunst	G1
MSTG 163.1 Gefährdung durch Sprengstoffe und giftige Gase ohne verbrech. Absicht	G1
MSTG 164.1 u. 2 u. 3 Herstellen, Verbergen, Weiterschaffen von Sprengstoffen und giftigen Gasen	G1
MSTG 62.1 Tötlichkeiten, Drohung	G1
MSTG 71 Tötlichkeiten, Drohung	G1
TSchG6 26.1 Tierquälerei	G1



<i>Gesetzesartikel</i>	<i>Gewichtung</i>
Schwerwiegende Verstösse gegen das SVG6 (diese Delikte gehören zur Kategorie allgemeine Delinquenz)	
SVG7 90.2 Grobe Verkehrsregelverletzung	
SVG 91.1 u. 2 Fahren in fahruntüchtigem Zustand	
SVG 95.1 u. 2 Fahren ohne Führerausweis	

Legende:

A StGB = Altes schweizerisches Strafgesetzbuch

StGB = Schweizerisches Strafgesetzbuch

SVG = Strassenverkehrsgesetz

AMSTG = Altes Militär-Strafgesetz

MSTG = Militär-Strafgesetz

TSchG = Tierschutzgesetz

Hinweise zur Gewichtung:

Es gibt einzelne wenige Delikte, bei denen im StGB nicht primär ein Strafrahmen sondern ein Mindeststrafmass angegeben ist. In diesen Fällen erfolgte die Gewichtung des jeweiligen Delikts analog vergleichbarer Delikte mit Angabe eines Strafmasses.

1 = Gewichtung: Delikte mit einem Strafrahmen bis max. 3 Jahre;

2 = Gewichtung: Delikte mit einem Strafrahmen bis max. 5 Jahre;

3 = Gewichtung: Delikte mit einem Strafrahmen von mehr als 5 Jahren.

G = Gewaltdelikt

S = Sexualdelikt

Quelle: Bewährungs- und Vollzugsdienste Zürich, Fachsupport und Lernprogramme (2013).



11 Anhang B: Beispiel einer FaST-Ausgabe

ROS

Risikoorientierter Sanktionenvollzug



Modellversuch B.J. der Kantone LU | SG | TG | ZH



© 2010 BVD Zürich
ROS CRIMS 2_01

Fallscreening (FaST)

Geschäfts-Nr.:

Personen-Nr.:

Klassifikation

A	Kein Klärungsbedarf	<input checked="" type="checkbox"/>
B	Bedarfsabklärung	<input type="checkbox"/>
C	Risikoabklärung + Bedarfsabklärung	<input type="checkbox"/>

SD / FaST:

Aktuellster Entscheid				
Datum	Behörde / Delikt	Strafmass	Vollzug	MN
Bemerkungen:				

Zusammenfassung			
Mit FaST wird die Indikation für eine Risiko- und/oder Bedarfsabklärung gestellt und keine Einschätzung des Delinquenzrisikos vorgenommen. Für eine B- oder C-Klassifikation müssen jeweils mehrere Kriterien zutreffen.			
Anlasstat	G	S	A
Deliktkategorien	0	0	1
Schweregrad / Strafmass			
Hohes deliktspezifisches Rückfallrisiko			
Massnahmen			
Vorstrafen	G	S	A
Deliktkategorien	0	0	0
Schweregrad / Strafmass			
Hohes deliktspezifisches Rückfallrisiko			
Massnahmen			
Mind. drei Vorstrafen			
Zusätzliche problematische Aspekte	zutreffend		
Mind. ein Gewaltdelikt im Alter unter 18 Jahren			
Mind. ein Sexualdelikt im Alter unter 18 Jahren			
Mind. ein Gewalt-/Sexualdelikt im Alter zwischen 18 und 25 Jahren			
Delikt im Zusammenhang mit häuslicher Gewalt			
Mind. zwei jugendanwaltschaftliche Einträge			
Mind. drei schwerwiegende Strassenverkehrsdelikte			
Widerrufe	<input checked="" type="checkbox"/>		
Polymorphe Delinquenz			
Risikominderung	zutreffend		
Mind. 5J «time at risk» seit Begehung des aktuellsten G-Delikts bzw. bei einer Freiheitsstrafe			
Mind. 5J «time at risk» im Rahmen von / nach aktuellster Massnahme nach G-/S-Delikt			
Aktuelles Alter mind. 50 Jahre			
Mind. 3 Vorstrafen			

12 Anhang C: Checkpunkte der ROS-Verlaufsliste „Therapie

1 Rechtliche Rahmenbedingungen der Therapie

Der Therapiebericht ...		ja	nein
1	nennt das Datum, an dem der Auftrag zur Therapie erteilt wurde (Tag, Monat, Jahr). (Siehe BF A1.1)		
2	nennt das Datum des strafrechtlichen Entscheids, der zum Vollzug führte. (Siehe BF A2.1)		
3	nennt den gesetzlichen Auftrag (Gesetzesartikel) und gibt Auskunft über das Vorliegen eines vorzeitigen Massnahmenantritts bzw. einer freiwilligen Therapie. (Siehe BF A2.3)		
4	nennt die gerichtlich und/oder behördlich angeordneten Weisungen/Auflagen. (Wenn keine angeordnet, "ja" ankreuzen) (Siehe BF A2.4)		

2 Setting und Grundlage der Therapie

- ☛ Prüfen Sie nur die Vollständigkeit der Angaben im Therapiebericht. Eine inhaltliche Prüfung ist nicht notwendig.

Der Therapiebericht ...		ja	nein
5	nennt die ROS-Abklärung als Informationsgrundlage für den Bericht. (Siehe BF B1.1)		
6	nennt das Datum des Therapiebeginns (Monat, Jahr). (Siehe BF B3.1)		
7	nennt den Zeitraum, auf den sich der Therapiebericht bezieht (Nennung von Anfang und Ende [Monat, Jahr]). (Siehe BF B3.2)		
8	nennt die Anzahl der Therapiesitzungen im Berichtszeitraum. (Siehe BF B3.3)		
9	nennt das Gesprächssetting, in welchem die Therapiesitzungen stattgefunden haben: Einzel- und/oder Gruppentherapie. (Siehe BF B3.3)		

3 Verlauf der Therapie und Beurteilung des Therapeuten

Teil A

- ☛ Prüfen Sie nur die Vollständigkeit der Angaben im Therapiebericht. Eine inhaltliche Prüfung ist nicht notwendig.

Der Therapiebericht ...		ja	nein
10	stellt dar, ob der Therapeut beim Klienten eine Diagnose nach ICD / DSM als vorliegend erachtet. (Siehe BF C1.2)		
11	beschreibt, dass alle im Fallkonzept (ROS-Abklärung) dargestellten Problembereiche für die Therapie übernommen wurden. (Siehe BF C2.3)		
12	enthält Aussagen über die Zuverlässigkeit des Klienten beim Einhalten der Termine. (Siehe BF D2.2)		
13	enthält Angaben über die Einhaltung von gerichtlich und/oder behördlich angeordneten Weisungen und/oder Auflagen. (Wenn keine angeordnet "ja" ankreuzen) (Siehe BF D2.3)		
14	enthält Aussagen über die im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF). (Siehe BF D4.1)		
15	nimmt Stellung zur Zweckmässigkeit der aktuellen Rahmenbedingungen der Therapie (Therapiesetting, -frequenz). (Siehe BF D6.1)		
16	weist das Rückfallrisiko des Klienten zu Beginn der Therapie gemäss mind. eines der nachfolgenden mechanischen Risk-Assessment Instrumente aus: FOTRES, VRAG, SORAG, Static-99, Static-2002, ODARA, DVRAG. (Siehe BF E1)		
17	beschreibt die aktuelle Einschätzung des Therapeuten bezüglich des Rückfallrisikos zum Berichtszeitpunkt. (Siehe BF E3)		

Teil B

- ☛ Teil B ist nur auszufüllen, falls eine medikamentöse Behandlung Bestandteil der deliktpräventiven Therapie ist.
- ☛ Falls dies der Fall ist, prüfen Sie nur die Vollständigkeit der Angaben im Therapiebericht. Eine inhaltliche Prüfung ist nicht notwendig.

Der Therapiebericht ...		ja	nein
18	enthält Aussagen über die Medikamenten-Compliance des Klienten. (Siehe BF D3.1a)		
19	enthält Aussagen über die Ansprechbarkeit des Klienten auf die Medikamente. (Siehe BF D3.1b)		

4 Hinweise auf einen kritischen Verlauf der Therapie

Teil A

☛ Prüfen Sie, ob der Therapiebericht Hinweise auf einen kritischen Verlauf der Therapie macht.

Der Therapiebericht ...		trifft nicht zu	trifft zu
20	verdeutlicht, dass weitere deliktrelevante Problembereiche erkannt wurden, die nicht im Fallkonzept (ROS-Abklärung) festgehalten sind. (Siehe BF C2.3)		
21	weist eine ungenügende Zuverlässigkeit des Klienten beim Einhalten von Terminen aus. (Siehe BF D2.2)		
22	weist aus, dass im Berichtszeitraum keine deliktpräventive Veränderung in den Problembereichen gemäss Fallkonzept (ROS-Abklärung) beobachtbar war. (Siehe BF D5.1)		
23	weist aus, dass die aktuellen Rahmenbedingungen der Therapie (Therapie setting, -frequenz) nicht oder nur fraglich geeignet sind, um dem Rückfallrisiko zweckmässig zu begegnen. (Siehe BF D6.1)		

Teil B


☛ Sollten Sie eine der nachfolgenden Aussagen als zutreffend bewerten, führt dies automatisch zu einem Eintrag in der Alarmliste und muss entsprechend bearbeitet werden.

Der Therapiebericht ...		trifft nicht zu	trifft zu
24	verdeutlicht, dass nicht alle deliktrelevanten Problembereiche gemäss ROS-Fallkonzept in den Fokus der Therapie gerückt sind. (Siehe BF D1.1)		
25	verdeutlicht, dass der Therapeut an keinem der deliktrelevanten Problembereiche gemäss Fallkonzept (ROS-Abklärung) gearbeitet hat. (Siehe BF D1.2)		
26	weist aus, dass gerichtlich und/oder behördlich angeordnete Weisungen und/oder Auflagen nicht eingehalten wurden. (Wenn keine angeordnet "trifft nicht zu" ankreuzen) (Siehe BF D2.3)		
27	weist eine ungenügende Compliance des Klienten in Bezug auf die Einnahme der zur Rückfallprävention wichtigen Medikamente aus. (Siehe BF D3.1a)		
28	weist auf eine kritische Entwicklung in den im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF) hin. (Siehe BF D4.2)		
29	verdeutlicht, dass die im Fallkonzept definierten deliktrelevanten äusseren Faktoren (LERF) nicht hinreichend überprüft wurden. (Siehe BF D4.2, wenn "unbekannt" oder nicht ausgefüllt, "trifft zu" ankreuzen)		




13 Anhang D: Merkblätter IR5 und IR6

Abbildung 8: Merkblatt für die Rater, exemplarisch für IR6



Kanton Zürich
Direktion der Justiz und des Innern
Amt für Justizvollzug
Bewährungs- und Vollzugsdienste
Projekt- und Qualitäts-Management



Universität
Zürich^{UZH}

Merkblatt

ROS-Verlaufsliste (VL), Interrater 6

Nicht vergessen:

- ✓ Arbeiten gemäss Instruktionen auf der ROS-Verlaufsliste
- ✓ Durchschnittlich 30 Minuten pro Fall
- ✓ Nicht mehr als 5 Fälle pro Tag bearbeiten
- ✓ Falllaufnummer und Kürzel in VL eintragen
- ✓ Wichtig: Mit anderen Ratern oder Kollegen und Kolleginnen weder über die Fälle selbst, noch über die Anwendung der Verlaufslisten sprechen

Agenda:

03.06.13	Kick-Off Veranstaltung, Übergabe der 30 Testfälle
28.06.13	Testlauf ist abgeschlossen
	Ausgefüllte VL und Kommentarliste per Mail an: damaris.reichen@ji.zh.ch
	Namenskonvention der zu sendenden Excel-Dateien:
	ROS-VL_IR6_[Fall-Laufnr.]_[KurzZ Rater]_[Datum (JJMMTT)].xls
	Beispiel: ROS-VL_IR6_1_MZ_130603.xls
Bis 05.07.13	Rückgabe der Testfälle (Berichtsformulare) an Damaris Reichen
Bis 03.07.13	Auswertung Testfälle durch UZH und Versand an ROS-Entwicklungsteam
04.07.13	Sitzung Austausch PPD-BVD-UZH: Resultate bewerten
Nach Absprache	Rückmeldung Auswertung Interrater-Test für Rater/innen

Kontakt:

Alex Schilling
ROS-Projektleitung,
Telefon 043 259 83 15
alex.schilling@ji.zh.ch

Mirjam Baur
ROS-Evaluation UZH
044 634 30 45
mirjam.baur@rwi.uzh.ch



14 Anhang E: Detailanalysen auf Item-Ebene, IR5

Item 1

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	2
22	1	1	1	2
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = -0.017

Item 2

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	2	2	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 0.322



Item 3

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	2	2	2	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	2	2	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	2	1	2	1
18	1	1	1	1
19	2	1	2	1
20	2	1	2	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 0.506

Item 4

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	2	2	2	2
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000



Item 5

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

Item 6

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000



Item 7

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = -0.008

Item 8

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000



Item 9

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	2	2	2	2
3	2	2	2	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	2	2	2	2
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	2	2	2	2
29	1	1	1	1
30	1	1	1	1

k = 1.000

Item 10

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	2
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = -0.008



Item 11

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	2	2	2	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	2	2	2
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	2	2	2	2
28	2	2	2	2
29	1	1	1	1
30	1	1	1	1

k = 0.924

Item 12

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	2	2	2
2	1	2	2	2
3	2	2	2	2
4	1	1	2	2
5	2	2	2	1
6	1	2	2	2
7	2	2	2	2
8	1	2	2	2
9	1	2	2	2
10	2	2	2	2
11	2	2	2	2
12	2	2	2	2
13	1	2	2	2
14	2	2	2	2
15	2	2	2	2
16	1	2	2	2
17	1	2	2	2
18	2	2	2	2
19	2	2	2	2
20	2	2	2	2
21	2	2	2	2
22	1	2	2	2
23	1	1	2	2
24	2	2	2	2
25	1	2	2	2
26	1	2	2	2
27	1	2	2	2
28	2	2	2	2
29	2	2	2	2
30	1	2	2	2

k = -0.089



Item 13

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	2	1	2	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	2

$k = 0.202$

Item 14

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	2
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	2	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

$k = -0.017$



Item 15

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	2
2	1	1	1	2
3	1	1	1	1
4	1	1	1	2
5	1	1	1	1
6	1	1	1	2
7	1	1	1	2
8	1	1	1	2
9	1	1	1	2
10	1	1	1	2
11	1	1	1	2
12	1	1	1	2
13	1	1	1	2
14	1	1	1	2
15	1	1	1	2
16	1	1	1	2
17	1	1	1	2
18	1	1	1	2
19	1	1	1	2
20	1	1	1	2
21	1	1	1	2
22	1	1	1	2
23	1	1	1	2
24	1	1	1	2
25	1	1	1	2
26	1	1	1	2
27	1	1	1	2
28	1	1	1	2
29	1	1	1	2
30	1	1	1	2

k = -0.304

Item 16

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000



Item 17

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	2
2	1	1	1	1
3	1	1	1	2
4	1	1	1	2
5	1	1	1	1
6	1	1	1	2
7	1		1	2
8	1	1	1	2
9	1	1	1	2
10	1	1	1	2
11	1	1	1	2
12	1	1	1	1
13	1	1	1	1
14	1	1	1	2
15	1	1	1	2
16	1	1	1	2
17	1	1	1	2
18	1	1	1	2
19	1	1	1	2
20	1	1	1	1
21	1	1	1	2
22	1	1	1	2
23	1	1	1	2
24	1	1	1	2
25	1	1	1	2
26	1	2	1	2
27	1	2	1	2
28	1	1	1	2
29	1	1	1	2
30	1	1	1	2

k = -0.228

Item 18

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	2	0	2	0
10	0	0	0	0
11	0	0	0	0
12	0	0	0	0
13	0	0	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	1	1	1	1
24	1	1	1	1
25	0	0	0	0
26	0	0	0	0
27	0	0	0	0
28	0	0	0	0
29	0	0	0	0
30	0	0	0	0

k = 0.857



Item 19

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	2	0	2	0
10	0	0	0	0
11	0	0	0	0
12	0	0	0	0
13	0	0	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	1	1	1	1
24	1	1	1	1
25	0	0	0	0
26	0	0	0	0
27	0	0	0	0
28	0	0	0	0
29	0	0	0	0
30	0	0	0	0

k = 0.857

Item 20

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	2	2	2	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	2	2	2
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	2	2	2	2
28	1	1	2	2
29	1	1	1	1
30	1	1	1	1

k = 0.799



Item 21

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	2	2	2	2
3	2	2	2	2
4	1	1	1	1
5	2	2	2	2
6	1	1	1	1
7	1	1	1	1
8	2	2	2	2
9	1	1	2	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	2	2	2
15	2	2	2	2
16	1	1	2	1
17	2	2	2	2
18	1	1	1	1
19	1	1	1	1
20	2	1	2	1
21	1	1	1	1
22	2	1	2	2
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	2	1	2	1
27	1	2	2	1
28	2	2	2	2
29	2	2	2	1
30	1	1	1	1

k = 0.718

Item 22

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	2	1
2	1	1	2	1
3	1	1	2	1
4	1	1	2	1
5	1	1	2	1
6	1	1	2	1
7	1	1	2	1
8	1	1	2	1
9	1	1	2	1
10	1	1	2	1
11	1	1	2	1
12	1	1	2	1
13	1	1	2	1
14	1	1	2	1
15	1	1	2	1
16	1	1	2	1
17	1	1	2	1
18	1	1	2	1
19	1	1	2	1
20	1	1	2	1
21	1	1	2	1
22	1	1	2	1
23	1	1	2	1
24	1	1	2	1
25	1	1	2	1
26	1	1	2	1
27	1	1	2	1
28	1	1	2	1
29	1	1	2	1
30	1	1	2	1

k = -0.333



Item 23

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	2	2	2	2
4	1	1	1	1
5	2	2	2	2
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

Item 24

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	2	1	2	1
7	1	2	2	2
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	1	2	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	2	1
26	1	1	1	1
27	2	2	2	2
28	1	1	2	1
29	1	1	1	1
30	1	1	1	1

k = 0.511



Item 25

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	2	1	2	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	2	1	2	2
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	2	2	2
28	1	1	2	1
29	2	1	2	2
30	1	1	1	1

k = 0.569

Item 26

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	2	1	1
3	1	2	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	2	2	1
16	1	1	1	1
17	1	1	2	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	2	2	2	2

k = 0.480



Item 27

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	2	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	2	1
9	1	1	2	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	2	2	2	2
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 0.557

Item 28

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	2	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	2	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = -0.017



Item 29

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	2
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	2	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = -0.020



15 Anhang F: Detailanalysen auf Item-Ebene, IR6

Item 1

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein

Item 2

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein



Item 3

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	2	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	2	1	1
16	1	1	1	1
17	1	1	1	1
18	1	2	1	1
19	1	1	1	1
20	1	1	1	1
21	1	2	2	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	2	1	1
27	1	2	1	1
28	1	2	1	1
29	1	2	1	1
30	1	1	1	1

k = -0.001

1 = ja; 2 = nein

Item 4

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	2	2	2	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	2	2	2	2
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	2	2	2	2
18	1	1	1	1
19	2	2	1	2
20	2	2	2	2
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 0.869

1 = ja; 2 = nein



Item 5

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	2	2	2	2
3	2	2	2	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	2	1	2
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	2	1	2
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	2	2	2	2
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	2	2	2	2
29	1	1	1	1
30	1	1	1	1

k = 0.840

1 = ja; 2 = nein

Item 6

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein



Item 7

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein

Item 8

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein



Item 9

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein

Item 10

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	2	1
6	1	1	1	1
7	1	1	2	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	2	1
13	1	1	1	1
14	1	1	2	1
15	1	1	1	1
16	1	1	2	2
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	2	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = -0.053

1 = ja; 2 = nein



Item 11

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	2	2	2	2
7	1	1	1	1
8	1	1	1	1
9	2	2	2	2
10	2	2	2	2
11	1	1	1	1
12	1	1	1	1
13	2	2	2	2
14	2	2	2	2
15	2	2	2	2
16	2	2	2	2
17	1	1	1	1
18	1	1	1	1
19	2	2	2	2
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	2	2	2	2
28	2	2	2	2
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein

Item 12

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein



Item 13

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	2	2	2	2
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein

Item 14

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein



Item 15

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein

Item 16

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = ja; 2 = nein



Item 17

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	2	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	2	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = -0.017

1 = ja; 2 = nein

Item 18

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	0	0	0	0
2	2	0	2	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	2	0	0	0
9	2	0	2	2
10	0	0	0	0
11	0	0	0	0
12	0	0	0	0
13	0	0	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	1	1	1	1
24	1	1	1	1
25	0	0	0	0
26	0	0	0	0
27	0	0	0	0
28	0	0	0	0
29	0	0	0	0
30	0	0	0	0

k = 0.739

1 = ja; 2 = nein



Item 19

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	0	0	0	0
2	2	0	2	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	2	0	0	0
9	2	0	2	2
10	0	0	0	0
11	0	0	0	0
12	0	0	0	0
13	0	0	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	1	1	1	1
24	1	1	1	1
25	0	0	0	0
26	0	0	0	0
27	0	0	0	0
28	0	0	0	0
29	0	0	0	0
30	0	0	0	0

k = 0.739

1 = ja; 2 = nein

Item 20

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	2	2	2	2
7	1	1	1	1
8	1	1	1	1
9	2	2	2	2
10	2	2	2	2
11	1	1	1	1
12	1	1	1	1
13	2	2	2	2
14	1	2	2	2
15	2	2	2	2
16	2	2	2	2
17	1	1	1	1
18	1	1	1	1
19	2	2	2	2
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	2	2	2
28	2	2	2	2
29	1	1	1	1
30	1	1	1	1

k = 0.923

1 = trifft nicht zu; 2 = trifft zu



Item 21

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	2	2	2
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	2	2	2	2
30	1	1	1	1

k = 1.000

1 = trifft nicht zu; 2 = trifft zu

Item 22

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	2	2	2
3	2	2	2	2
4	1	1	1	1
5	2	2	2	2
6	1	1	1	1
7	1	1	1	1
8	2	2	2	2
9	2	2	2	2
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	2	2	2
15	2	2	2	2
16	2	2	2	2
17	2	2	2	2
18	1	1	1	1
19	1	1	1	1
20	1	2	2	2
21	1	1	1	1
22	2	2	2	2
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	2	2	2	2
27	2	2	2	2
28	2	2	2	2
29	2	2	2	2
30	1	1	1	1

k = 0.933

1 = trifft nicht zu; 2 = trifft zu



Item 23

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	2	2	2	2
4	1	1	1	1
5	2	2	2	2
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 1.000

1 = trifft nicht zu; 2 = trifft zu

Item 24

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	2	2	2	2
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	2	2	2
26	1	1	1	1
27	2	2	2	2
28	1	1	1	2
29	1	1	1	1
30	1	1	1	1

k = 0.815

1 = trifft nicht zu; 2 = trifft zu



Item 25

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	2	2	2	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	2	2	2	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	2	1
15	1	1	1	1
16	1	2	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	2	2	2
28	1	1	1	2
29	1	2	2	1
30	1	1	1	1

k = 0.518

1 = trifft nicht zu; 2 = trifft zu

Item 26

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	2	2	2	1
16	1	1	1	1
17	2	2	2	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	2	2	2	2

k = 0.782

1 = trifft nicht zu; 2 = trifft zu



Item 27

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	2	2
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	2	2	2	2
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 0.766

1 = trifft nicht zu; 2 = trifft zu

Item 28

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	2	2	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 0.322

1 = trifft nicht zu; 2 = trifft zu



Item 29

Fall-Nr.	Rater 1	Rater 2	Rater 3	Rater 4
1	1	1	1	2
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	2	2	2	2
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

k = 0.791

1 = trifft nicht zu; 2 = trifft zu