

**Daniel Wolf
Sarah Schneider
Astrid Schütz**

SEC-I und SEC-SJT

**Inventar und Situational Judgement Test
zur Erfassung sozial-emotionaler Kompetenz(en)**

Testmanual



Kofinanziert durch das
Programm Erasmus+
der Europäischen Union

Daniel Wolf, Sarah Schneider und Astrid Schütz

SEC-I und SEC-SJT

Inventar und Situational Judgement Test
zur Erfassung sozial-emotionaler Kompetenz(en)

Diagnostische Verfahren
zur multimethodalen Erfassung sozial-emotionaler Kompetenz(en)
bei jungen Menschen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung

Testmanual

Autor und Autorinnen

Daniel Wolf, M.Sc., und **Sarah Schneider, M.Sc.**, sind wissenschaftlicher Mitarbeiter und wissenschaftliche Mitarbeiterin am Lehrstuhl für Persönlichkeitspsychologie und Psychologische Diagnostik der Otto-Friedrich-Universität Bamberg.

Prof. Dr. Astrid Schütz, ist Inhaberin des Lehrstuhls für Persönlichkeitspsychologie und Psychologische Diagnostik der Otto-Friedrich-Universität Bamberg und Leiterin des Kompetenzzentrums für Angewandte Personalpsychologie der Otto-Friedrich-Universität Bamberg.

Für die Mitarbeit danken der Autor und die Autorinnen Mia Degro; Katharina Ferber; Annette Göller, B.Sc.; Magdalena Höß, B.Sc.; Angelika Münch; Daniela Reichelt, M.Sc.; Laura Stolle; Madeleine Welker und Gesine Winter.

Kontakt

Prof. Dr. Astrid Schütz
Lehrstuhl für Persönlichkeitspsychologie und Psychologische Diagnostik
Otto-Friedrich-Universität Bamberg
96047 Bamberg
E-Mail: astrid.schuetz@uni-bamberg.de

Die Homepage des Lehrstuhls für Persönlichkeitspsychologie und Psychologische Diagnostik der Otto-Friedrich-Universität Bamberg erreichen Sie unter <https://www.uni-bamberg.de/perspsych/>.

Lizenz

Testmanual, Testmaterialien und Auswertungsdatei sind Ergebnisse des von der Europäischen Kommission geförderten Erasmus+-Projekts SEC4VET (Assessing and Promoting Social-Emotional Competences of Young People with Cognitive Impairments in Vocational Education and Training, Förderkennzeichen 2017-1-DE02-KA202-004136; die Homepage des Projekts erreichen Sie unter <https://sec4vet.eu/>) und werden der Öffentlichkeit im Rahmen einer Creative-Commons-Lizenz CC BY-SA 4.0 (<http://creativecommons.org/licenses/by-sa/4.0/>) kostenfrei zur Verfügung gestellt.



Haftungsausschluss

Die Unterstützung der Europäischen Kommission für die Erstellung dieser Veröffentlichung stellt keine Billigung des Inhalts dar, welcher nur die Ansichten der Verfasser wiedergibt, und die Kommission kann nicht für eine etwaige Verwendung der darin enthaltenen Informationen haftbar gemacht werden.

Sämtliche Inhalte des Testmanuals sowie Testmaterialien und Auswertungsdatei wurden mehrfach sorgfältig geprüft und Korrektur geleistet. Trotz strenger Qualitätskontrollen ist es nicht möglich, Fehler vollständig auszuschließen. Autor und Autorinnen können für mögliche Folgen keine Haftung übernehmen. Für Rückmeldungen und Anregungen wenden Sie sich bitte direkt an Frau Prof. Dr. Astrid Schütz (siehe Kontakt).

Bibliografische Angaben

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar. <https://nbn-resolving.org/urn:nbn:de:bvb:473-irb-48863>

1. Auflage 2020

<https://doi.org/10.20378/irb-48863>

Zitation

Wolf, D., Schneider, S., & Schütz, A. (2020). *Inventar und Situational Judgement Test zur Erfassung sozial-emotionaler Kompetenz(en) (SEC-I und SEC-SJT)*. Otto-Friedrich-Universität Bamberg. <https://doi.org/10.20378/irb-48863>

Inhalt

Vorwort	7
Kurzbeschreibung	8
1. Übersicht	11
1.1. Überblick über Inventar und Situational Judgement Test	12
1.1.1. Inventar zur Erfassung sozial-emotionaler Kompetenz(en)	12
1.1.2. Situational Judgement Test zur Erfassung sozial-emotionaler Kompetenz	15
1.2. Testmaterialien	15
1.3. Aufbau des Testmanuals	16
1.4. Diagnostische Fragestellung	17
1.5. Theoretischer Hintergrund	18
1.6. Konstruktion von Inventar und Situational Judgement Test	21
1.7. Workshop zur Diagnostik sozial-emotionaler Kompetenzen	25
2. Anwendung	27
2.1. Anwendungsbereich	28
2.2. Durchführung	29
2.2.1. Qualifikation der Testleitung	29
2.2.2. Anforderungen an die Testperson	30
2.2.3. Allgemeine Hinweise zur Erhebung	30
2.2.4. Gestaltung der Erhebungssituation	31
2.2.5. Instruktion der Testperson(en)	32
2.2.6. Umgang mit Fragen und Störungen während der Erhebung	33
2.3. Auswertung	34
2.3.1. Qualifikation der auswertenden Person	34
2.3.2. Allgemeine Hinweise zur Auswertung	34
2.3.3. Prüfung der Qualität der Bearbeitung	35
2.3.4. Auswertung von Hand	40
2.3.5. Auswertung mittels Auswertungsbögen	43
2.3.6. Computerbasierte Auswertung	50

2.4.	Interpretation	50
2.4.1.	Qualifikation der interpretierenden Person	50
2.4.2.	Allgemeine Hinweise zur Interpretation	51
2.4.3.	Grundlagen der Interpretation psychologischer Testergebnisse	53
2.4.4.	Transformation von Rohwerten in standardisierte Werte	59
2.4.5.	Interpretation einzelner Skalen	60
2.4.6.	Interpretation einer Profillinie	61
2.4.7.	Interpretation von Differenzen	63
2.4.8.	Inhaltliche Beschreibung der Skalen	68
2.4.9.	Situational Judgement Test als Screeningverfahren	73
2.4.10.	Ergebnisrückmeldungen	75
2.4.11.	Interpretation mit dem zugehörigen Profilbogen	76
2.4.12.	Interpretationsbeispiele	81
3.	Testtheoretische Fundierung	89
3.1.	Psychometrische Merkmale	90
3.1.1.	Itemanalyse	90
3.1.2.	Skalenanalyse	92
3.1.3.	Skaleninterkorrelationen	93
3.2.	Testgüte	93
3.2.1.	Objektivität	93
3.2.2.	Reliabilität	94
3.2.3.	Validität	97
4.	Normierung	119
4.1.	Rekrutierung der Gesamtstichprobe	120
4.2.	Beschreibung der Gesamtstichprobe	120
4.2.1.	Alter und Geschlecht	121
4.2.2.	Bildungsstand und Ausbildungsstatus	121
4.2.3.	Störungsbilder und Behinderungen	121
4.3.	Berechnung der Normwerte	122
4.3.1.	Verteilung der Skalenrohwerte	122
4.3.2.	Normierung der Rohwerte	123
4.3.3.	Normen und Subnormen	124
	Literatur	129

Anhang.....	135
Tabelle A1	136
Tabelle A2	138
Tabelle A3	140
Tabelle A4	142
Tabelle A5	144
Tabelle A6	146
Tabelle A7	148
Tabelle A8	149
Tabelle A9	150
Tabelle A10	151
Tabelle A11	152
Tabelle A12	153
Tabelle A13	154
Tabelle A14	155
Tabelle A15	157
Tabelle A16	159
Tabelle A17	160
Tabelle A18	161
Abbildung A1	162
Abbildung A2	163
Abbildung A3	164

Vorwort

Sozial-emotionale Kompetenzen werden häufig als Schlüsselqualifikationen angesehen und sind in vielen Berufen unerlässlich. Daher ist ihrer Förderung auch im Rahmen der beruflichen Inklusion eine bedeutende Rolle beizumessen. Doch verfügbare Testverfahren sind in der Regel stark kognitiv orientiert – wodurch sich die standardisierte Erfassung und gezielte Förderung sozial-emotionaler Kompetenzen bei jungen Menschen mit kognitiven Beeinträchtigungen schwierig gestaltet. Angetrieben von der in Berufsbildungswerken wiederkehrenden Beobachtung, dass bestimmte Personen trotz kognitiver Beeinträchtigung hervorragend in der Lage sind, sich in andere einzufühlen und bei Bedarf Trost zu spenden, entstand vor einigen Jahren in Gesprächen mit Manfred Weiser, Berufsbildungswerk Mosbach-Heidelberg (Mosbach und Heidelberg), die Idee, sozial-emotionale Kompetenzen möglichst unabhängig von kognitiven Fähigkeiten zu erfassen. Basierend auf dieser Idee und koordiniert durch Martin Künemund, Berufsbildungswerk Josefsheim (Bigge), entstand ein mehrjähriges internationales Projekt mit dem Anliegen, sozial-emotionale Kompetenzen bei jungen Menschen mit kognitiven Beeinträchtigungen standardisiert zu erfassen und gezielt zu fördern. Im Rahmen dieses Projekts wurden die vorliegenden diagnostischen Verfahren zur multimethodalen Erfassung sozial-emotionaler Kompetenz(en) bei jungen Menschen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelt. Mit diesem Testmanual liegt eine umfassende Dokumentation dieser, für die Zielgruppe in leichter Sprache konzipierten, wissenschaftlich fundierten und praktisch erprobten Verfahren vor. Der Fragebogen zur Selbsteinschätzung, der Fragebogen zur Fremdeinschätzung und der Situational Judgement Test können zur Erfüllung dieses Anliegens gemeinsam, aber auch unabhängig voneinander eingesetzt werden. Wir hoffen, dass die kostenfrei zur Verfügung gestellten diagnostischen Verfahren dazu beitragen werden, die berufliche Inklusion von jungen Menschen mit Beeinträchtigungen weiter voranzutreiben.

Ein besonderer Dank gebührt an dieser Stelle allen an der Entwicklung und Erprobung der diagnostischen Verfahren aktiv beteiligten Personen sowie den vielen jungen Menschen, die die diagnostischen Verfahren freiwillig in verschiedensten Stadien der Entwicklung bearbeitet haben. Weiterer Dank gebührt der Europäischen Kommission für die finanzielle Förderung des Projekts, ohne die ein so umfangreiches Vorhaben nicht hätte realisiert werden können.

Bamberg, im Herbst 2020

Astrid Schütz, Sarah Schneider und Daniel Wolf

Kurzbeschreibung

Anwendungsbereich

Jugendliche und junge Erwachsene mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren. Anwendung insbesondere in der Förderung von Menschen mit Lernbehinderungen.

Erfasste Merkmale

Sozial-emotionale Kompetenz sowie acht spezifische Subfacetten aus den Bereichen Empathie (Empathisches Verhalten, Kognitive Empathie), Umgang mit Emotionen (Wahrnehmung der Emotionen anderer, Wahrnehmung eigener Emotionen, Regulation der Emotionen anderer, Regulation eigener Emotionen) und prosozialer Orientierung (Positive Beziehungsgestaltung, Beachtung sozialer Normen).

Besonderheiten

Aufeinander abgestimmte multimethodale Erfassung sozial-emotionaler Kompetenz(en). Subjektiver (Selbstbild) wie auch mehrstufig objektivierbarer Zugang (Fremdbild aus Perspektive einer oder mehrerer anderer Personen) zu erfassten Merkmalen. Inventar und Situational Judgement Test in leichter Sprache. Vermeidung von Überforderung. Fragebogen zur Fremdeinschätzung mit eigenständiger Norm. Auf praktische Bedürfnisse zugeschnitten, praktisch erprobt und wissenschaftlich fundiert.

Testmaterialien

Testmanual, Fragebogen zur Selbsteinschätzung (SEC-I-SR), Fragebogen zur Fremdeinschätzung (SEC-I-OR), Situational Judgement Test (SEC-SJT), SEC-I-SR Auswertungsbogen, SEC-I-OR Auswertungsbogen, SEC-SJT Auswertungsbogen, Profilbogen, Auswertungsdatei (Microsoft® Excel® Arbeitsmappe).

Anwendung

Detaillierte Hinweise zur Durchführung, Auswertung und Interpretation. Durchführung ohne Zeitlimit, Bearbeitungsdauer des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung ca. 10 Minuten, Bearbeitungsdauer des Situational Judgement Tests ca. 30 Minuten. Computerbasierte Auswertung und Interpretation durch erfahrene Person in ca. 10 Minuten.

Objektivität

Durch Standardisierung von Durchführung, Auswertung und Interpretation sichergestellt.

Reliabilität

Interne Konsistenz anhand Cronbachs Alpha, für Globalskala des Fragebogens zur Selbsteinschätzung $\alpha = .92$, Globalskala des Fragebogens zur Fremdeinschätzung $\alpha = .96$, Globalskala des Situational Judgement Test $\alpha = .65$, Subskalen des Fragebogens zur Selbsteinschätzung $.68 \leq \alpha \leq .82$, $M = .76$ und Subskalen des Fragebogens zur Fremdeinschätzung $.81 \leq \alpha \leq .91$, $M = .86$.

Validität

Empirische Untersuchung der diagnostischen Verfahren in einer Gesamtstichprobe von $N = 306$ Personen. Konstruktvalidität anhand von Hauptkomponentenanalysen und Multitrait-Multimethod-Analyse. Ergänzend divergente Validität bezüglich Intelligenz anhand der Standard Progressive Matrices in einer Substichprobe von $n = 36$ (Alter: $M = 18.97$, $s = 2.06$, Geschlecht: 13.9% weiblich). Kriteriumsvalidität anhand von Störungsbildern gemäß ICD-10 Kapitel V, des Bildungsstands und des Ausbildungsstatus.

Normierung

Normierung der diagnostischen Verfahren an einer Stichprobe von $N = 306$ Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15 bis 24 Jahren ($MD = 17$, $M = 18.72$, $s = 2.32$) weiblichen (36.3%) und männlichen (63.7%) Geschlechts bzw. deren Kontaktpersonen, denen die einzuschätzenden Jugendlichen und jungen Erwachsenen seit mindestens einem Monat ($MD = 6$) bekannt waren. Ergänzende Subnormen für weibliches ($n = 111$, Alter: $M = 18.86$, $s = 2.51$) und männliches Geschlecht ($n = 195$, Alter: $M = 18.64$, $s = 2.21$). Für den Fragebogen zur Selbsteinschätzung und den Fragebogen zur Fremdeinschätzung liegen T -Werte und Prozentränge, für den Situational Judgement Test Prozentränge vor.

Sprache

Deutschsprachige Originalversion.

1. Übersicht

Das erste Kapitel gibt eine allgemeine Übersicht über die diagnostischen Verfahren und die zugehörigen Testmaterialien. Der Aufbau des Testmanuals wird beschrieben um die Arbeit mit diesem zu erleichtern. Um sicherzustellen, dass die Verfahren möglichst korrekt eingesetzt werden, werden die diagnostische Fragestellung und der theoretische Hintergrund ausführlich besprochen. Darüber hinaus wird die Konstruktion der diagnostischen Verfahren beschrieben.

1.1. Überblick über Inventar und Situational Judgement Test

Das Inventar zur Erfassung sozial-emotionaler Kompetenz(en) (SEC-I, Abkürzung für engl. Social-Emotional Competences Inventory) und der Situational Judgement Test zur Erfassung sozial-emotionaler Kompetenz (SEC-SJT, Abkürzung für engl. Social-Emotional Competence Situational Judgement Test) wurden für die standardisierte multimethodale Erfassung sozial-emotionaler Kompetenz(en) bei jungen Menschen mit subklinischer oder klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelt. Abbildung 1 (auf der nächsten Seite) zeigt eine Gesamtübersicht der mit den diagnostischen Verfahren multimethodal erfassbaren Merkmale.

Die diagnostischen Verfahren sind für die multimethodale Erfassung des aktuellen Status sozial-emotionaler Kompetenz(en) speziell aufeinander abgestimmt und entfalten ihr volles Potential, wenn alle drei Verfahren bei einer Testperson eingesetzt werden und die Ergebnisse kombiniert bzw. integriert werden. Der Fragebogen zur Selbsteinschätzung, der Fragebogen zur Fremdeinschätzung und der Situational Judgement Test können aber auch unabhängig voneinander eingesetzt werden. Instruktion und Items der diagnostischen Verfahren sind in leichter Sprache verfasst.

1.1.1. Inventar zur Erfassung sozial-emotionaler Kompetenz(en)

Das Inventar (SEC-I) ermöglicht sowohl die Erfassung globaler sozial-emotionaler Kompetenz wie auch die differenzierte Erfassung spezifischer Subfacetten sozial-emotionaler Kompetenzen (aus den Bereichen Empathie, Umgang mit Emotionen und prosozialer Orientierung) bei jungen Menschen mit subklinischer oder klinischer kognitiver bzw. psychischer Beeinträchtigung. Die Subfacetten ermöglichen gegenüber der globalen sozial-emotionalen Kompetenz eine differenziertere Betrachtung verschiedener Aspekte sozial-emotionaler Kompetenzen (siehe Abbildung 1).

Abbildung 1

Übersicht über die diagnostischen Verfahren und die multimethodal erfassbaren Merkmale

SEC-I Inventar zur Erfassung sozial-emotionaler Kompetenz(en)		SEC-SJT Situational Judgement Test zur Erfassung sozial-emotionaler Kompetenz
SEC-I-SR Fragebogen zur Selbsteinschätzung	SEC-I-OR Fragebogen zur Fremdeinschätzung	
Globalskala SEC-SR Sozial-emotionale Kompetenz (Selbstbild)	Globalskala SEC-OR Sozial-emotionale Kompetenz (Fremdbild)	Globalskala SEC-SJT Sozial-emotionale Kompetenz (Leistung)
Subskala EBE-SR Empathisches Verhalten (Selbstbild) Subskala ECO-SR Kognitive Empathie (Selbstbild) Subskala PEA-SR Wahrnehmung der Emotionen anderer (Selbstbild) Subskala PES-SR Wahrnehmung eigener Emotionen (Selbstbild) Subskala REA-SR Regulation der Emotionen anderer (Selbstbild) Subskala RES-SR Regulation eigener Emotionen (Selbstbild) Subskala SOB-SR Positive Beziehungsgestaltung (Selbstbild) Subskala SON-SR Beachtung sozialer Normen (Selbstbild)	Subskala EBE-OR Empathisches Verhalten (Fremdbild) Subskala ECO-OR Kognitive Empathie (Fremdbild) Subskala PEA-OR Wahrnehmung der Emotionen anderer (Fremdbild) Subskala PES-OR Wahrnehmung eigener Emotionen (Fremdbild) Subskala REA-OR Regulation der Emotionen anderer (Fremdbild) Subskala RES-OR Regulation eigener Emotionen (Fremdbild) Subskala SOB-OR Positive Beziehungsgestaltung (Fremdbild) Subskala SON-OR Beachtung sozialer Normen (Fremdbild)	

Anmerkung. Im Gegensatz zum Fragebogen zur Selbsteinschätzung (SEC-I-SR) und zum Fragebogen zur Fremdeinschätzung (SEC-I-OR) erfasst der Situational Judgement Test (SEC-SJT) keine Subskalen.

Das Inventar wurde in deutscher Sprache entwickelt, validiert und normiert. Es liegt ein Fragebogen zur Selbsteinschätzung und ein Fragebogen zur Fremdeinschätzung vor. Alle Items sind in leichter Sprache formuliert und auf einer vollständig verbal verankerten fünf-stufigen Likert-Skala zu beantworten.

1.1.1.1. Fragebogen zur Selbsteinschätzung

Der Fragebogen zur Selbsteinschätzung (SEC-I-SR) erfasst mit insgesamt 41 Items globale sozial-emotionale Kompetenz auf der Globalskala Sozial-emotionale Kompetenz (SEC-SR, Abkürzung für engl. Social-Emotional Competence) mit allen 41 Items, sowie Subfacetten sozial-emotionaler Kompetenz auf den acht Subskalen:

- ◆ Empathisches Verhalten (EBE-SR, Abkürzung für engl. Empathic Behavior) mit 4 Items,

- ◆ Kognitive Empathie (ECO-SR, Abkürzung für engl. Cognitive Empathy) mit 5 Items,
- ◆ Wahrnehmung der Emotionen anderer (PEA-SR, Abkürzung für engl. Perception of Emotions in Others) mit 5 Items,
- ◆ Wahrnehmung eigener Emotionen (PES-SR, Abkürzung für engl. Perception of Emotions in Oneself) mit 5 Items,
- ◆ Regulation der Emotionen anderer (REA-SR, Abkürzung für engl. Regulation of Emotions in Others) mit 5 Items,
- ◆ Regulation eigener Emotionen (RES-SR, Abkürzung für engl. Regulation of Emotions in Oneself) mit 4 Items,
- ◆ Positive Beziehungsgestaltung (SOB-SR, Abkürzung für engl. Socially Oriented Behavior) mit 9 Items und
- ◆ Beachtung sozialer Normen (SON-SR, Abkürzung für engl. Respect for Social Norms) mit 4 Items.

Weitere Informationen zu den Inhalten der Skalen finden sich im Abschnitt *Inhaltliche Beschreibung der Skalen* (siehe S. 68). Mit dem Fragebogen zur Selbsteinschätzung wird die subjektive Perspektive der Testperson erfasst. Einschätzungen der eigenen Kompetenz entsprechen nicht notwendigerweise Fremdeinschätzungen oder Ergebnissen aus Leistungstests (siehe Abschnitt *Theoretischer Hintergrund*, S. 18; und Abschnitt *Verzerrungen im Antwortverhalten*, S. 36). Durch die Erfassung der subjektiven Perspektive wird allerdings ein Aspekt inneren emotionalen Erlebens zugänglich, der in Kombination mit einer bzw. weiteren Fremdeinschätzung(en) und dem Ergebnis aus dem Situational Judgement Test eine ganzheitliche Sicht auf die Testperson vervollständigt.

1.1.1.2. Fragebogen zur Fremdeinschätzung

Der Fragebogen zur Fremdeinschätzung (SEC-I-OR) ist parallel zum Fragebogen zur Selbsteinschätzung konstruiert. Alle Items sind in der dritten (sie/er) anstelle der ersten Person Singular (ich) formuliert. Aufgrund der geschlechtsneutralen Formulierung kann der Fragebogen zur Fremdeinschätzung für Personen aller Geschlechter eingesetzt werden.

Parallel zum Fragebogen zur Selbsteinschätzung erfasst der Fragebogen zur Fremdeinschätzung globale sozial-emotionale Kompetenz auf der inhaltsgleichen Globalskala Sozial-emotionale Kompetenz (SEC-OR) sowie die acht Subfacetten sozial-emotionaler Kompetenz auf den inhaltsgleichen Subskalen (für den Fragebogen zur Fremdeinschätzung in der Abkürzung mit der Endung -OR anstelle der Endung -SR bezeichnet). Weitere Infor-

mationen über die zum Fragebogen zur Selbsteinschätzung inhaltsgleichen Skalen des Fragebogens zur Fremdeinschätzung sind im Abschnitt *Fragebogen zur Selbsteinschätzung* (siehe S. 13) sowie im Abschnitt *Inhaltliche Beschreibung der Skalen* (siehe S. 68) zu finden.

Mit dem Fragebogen zur Fremdeinschätzung kann die subjektive Perspektive der eigenen Person um die objektivierte Perspektive aus Sicht einer oder mehrerer anderer Personen erweitert werden. Ferner kann der Fragebogen zur Fremdeinschätzung auch dann eingesetzt werden, wenn eine Selbsteinschätzung nicht möglich ist oder Zweifel an der Qualität der Bearbeitung des Fragebogens zur Selbsteinschätzung bestehen.

1.1.2. Situational Judgement Test zur Erfassung sozial-emotionaler Kompetenz

Mit dem Situational Judgement Test (SEC-SJT) wurde ein (Kurz-)Leistungstests zur objektiven Erfassung globaler sozial-emotionaler Kompetenz bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung konstruiert. Das Potential (Wissen bzw. Erfahrung), sich in verschiedensten Situationen sozial-emotional kompetent verhalten zu können, wird mit 10 Items auf einer einzelnen Skala (SEC-SJT) erfasst. Zugunsten der Zumutbarkeit wurden mit nur 10 Items Abstriche im Hinblick auf eine differenzierte Erfassung von Subfacetten sowie die Reliabilität in Kauf genommen. Weitere Informationen zum Inhalt der Globalskala SEC-SJT finden sich im Abschnitt *Inhaltliche Beschreibung der Skalen* (siehe S. 68).

Mit dem Situational Judgement Test können die Ergebnisse aus subjektiver Selbsteinschätzungsperspektive und Fremdeinschätzungsperspektive(n) um eine objektive Perspektive erweitert werden. Als Screening eingesetzt kann der Situational Judgement Test Hinweise auf potentielle Defizite geben, sodass die Erhebung des Fragebogens zur Selbsteinschätzung bzw. des Fragebogens zur Fremdeinschätzung in Erwägung gezogen werden sollte (siehe Abschnitt *Situational Judgement Test als Screeningverfahren*, S. 73).

1.2. Testmaterialien

Die zur Verfügung gestellten Materialien umfassen:

- ◆ dieses Testmanual, das Anwendung, testtheoretische Fundierung sowie Normierung der diagnostischen Verfahren umfassend dokumentiert,
- ◆ den Fragebogen zur Selbsteinschätzung (SEC-I-SR) mit pronominaler Anrede in der zweiten Person Singular (du), die Items sind in der ersten Person Singular (ich) formuliert,

- ◆ den Fragebogen zur Fremdeinschätzung (SEC-I-OR) mit pronominaler Anrede in der dritten Person Plural (sie), die Items sind geschlechtsneutral in der dritten Person Singular (sie/er) formuliert,
- ◆ den Situational Judgement Test (SEC-SJT) mit pronominaler Anrede in der zweiten Person Singular (du),
- ◆ den SEC-I-SR Auswertungsbogen, für die Auswertung des Fragebogens zur Selbsteinschätzung,
- ◆ den SEC-I-OR Auswertungsbogen, für die Auswertung des Fragebogens zur Fremdeinschätzung,
- ◆ den SEC-SJT Auswertungsbogen, für die Auswertung des Situational Judgement Tests,
- ◆ den Profilbogen, in dem die Ergebnisse aus dem Fragebogen zur Selbsteinschätzung, dem Fragebogen zur Fremdeinschätzung und dem Situational Judgement Test gemeinsam dokumentiert werden sowie
- ◆ eine Auswertungsdatei (Microsoft® Excel® Arbeitsmappe) für die gemeinsame computerbasierte Auswertung des Fragebogens zur Selbsteinschätzung, des Fragebogens zur Fremdeinschätzung und des Situational Judgement Tests.

1.3. Aufbau des Testmanuals

Im Wesentlichen enthält dieses Testmanual einen praktischen (Kapitel *Anwendung*, ab S. 27) und einen empirischen Teil (Kapitel *Testtheoretische Fundierung*, ab S. 89). Ergänzend eine Einführung (Kapitel *Übersicht*, ab S. 11) sowie detaillierte Informationen zur Normierung (Kapitel *Normierung*, ab S. 119).

Der praktische Teil (Kapitel *Anwendung*, ab S. 27) ist dem empirischen Teil vorangestellt und richtet sich vor allem an diejenigen, die mit dem Fragebogen zur Selbsteinschätzung, dem Fragebogen zur Fremdeinschätzung wie auch dem Situational Judgement Test arbeiten. Das Kapitel beinhaltet die Beschreibung des Anwendungsbereichs sowie technische Hinweise zur Durchführung, Auswertung und Interpretation. Werden Durchführung und Auswertung von psychologischen Hilfskräften übernommen, sollten diese mit den entsprechenden Abschnitten dieses Kapitels vertraut sein. Der Abschnitt zur Interpretation schneidet wichtige *Grundlagen der Interpretation psychologischer Testergebnisse* (siehe S. 53) an und erläutert verschiedene methodische Hilfsmittel (beispielsweise Vertrauensbereiche und kritische Differenzen). Die Interpretation von Testergebnissen sollte jedoch nur durch Fachpersonal mit psychologischer Ausbildung und Erfahrung in psychologischer Diagnostik erfolgen. Für eine angemessene Interpretation ist es unerlässlich, sich mit den Inhalten

des empirischen Teils vertraut zu machen (Kapitel *Testtheoretische Fundierung*, ab S. 89). Dieser beinhaltet die Ergebnisse der empirischen Überprüfung von Inventar und Situational Judgement Test: die Untersuchung psychometrischer Merkmale (siehe Abschnitt *Psychometrische Merkmale*, S. 90) und der Testgüte (siehe Abschnitt *Testgüte*, S. 93) sowie deren Bewertung.

Querverweise zu anderen, für das Verständnis des jeweiligen Abschnitts relevanten Abschnitten, finden sich mit Angabe der entsprechenden Seitenzahl direkt im Text. Um wichtige methodische Hilfsmittel und empirische Befunde übersichtlich zu bündeln, befindet sich ein Großteil der Tabellen im Anhang. Verweise auf die jeweils relevanten Tabellen des Anhangs finden sich direkt im Text.

1.4. Diagnostische Fragestellung

Der Fragebogen zur Selbsteinschätzung, der Fragebogen zur Fremdeinschätzung und der Situational Judgement Test wurden speziell für die Erfassung von Förderbedarf bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Bereich sozial-emotionaler Kompetenz(en) entwickelt. Die Bestimmung des aktuellen Status ermöglicht die bedarfsgerechte Auswahl und Planung von Maßnahmen zur gezielten individuellen Förderung.

Die diagnostischen Verfahren wurden für die multimethodale Erfassung des aktuellen Status sozial-emotionaler Kompetenz(en) von jungen Menschen mit subklinischer oder klinischer kognitiver bzw. psychischer Beeinträchtigung aufeinander abgestimmt. Mit der multimethodalen Erfassung können Perspektiven sowohl mit verschiedenen Methoden als auch aus verschiedenen Informationsquellen erschlossen und zu einer Gesamtperspektive integriert werden. Die multimethodale Erfassung liefert einen subjektiven wie auch mehrstufig objektivierbaren Zugang zu sozial-emotionaler Kompetenz. Durch den Einbezug von mehr als einer Fremdeinschätzung für eine Testperson, kann die Fremdeinschätzung zunehmend objektiviert werden.

Die drei diagnostischen Verfahren entfalten ihr Potential insbesondere dann, wenn alle drei Verfahren bei einer Testperson ergänzend zum Einsatz kommen. Es ist aber auch möglich, nur ein Verfahren einzusetzen. Bestehen beispielsweise Zweifel an der Qualität der Bearbeitung des Fragebogens zur Selbsteinschätzung kann der Fragebogen zur Fremdeinschätzung oder der Situational Judgement Test eingesetzt werden.

Nur der kombinierte Einsatz ermöglicht die Betrachtung von Konvergenzen und Divergenzen in subjektiver (Selbsteinschätzung), mehrstufig objektivierbarer (Fremdeinschätzung) und objektiver Perspektive (Leistung im Situational Judgement Test). Alle Verfahren er-

fassen globale sozial-emotionale Kompetenz. Zusätzlich ermöglicht das Inventar die Differenzierung globaler sozial-emotionaler Kompetenz in acht spezifische Subfacetten aus den Bereichen Empathie, Umgang mit Emotionen und prosoziale Orientierung.

Die differenzierte Erfassung sozial-emotionaler Kompetenz erlaubt die Identifikation von Defiziten und Ressourcen. Während Defizite die Indikation für eine Intervention liefern, zeigen Ressourcen Anknüpfungspunkte, an denen eine Intervention ansetzen kann. Mit der differenzierten Erfassung spezifischer Subfacetten sozial-emotionaler Kompetenz aus den Bereichen Empathie, Umgang mit Emotionen und prosozialer Orientierung können spezifische Ressourcen identifiziert werden, an die zum Abbau eines Defizits sozial-emotionaler Kompetenz angeknüpft werden kann. Aufgrund der Homogenität der Subfacetten bzw. der relativ hohen Interkorrelation der Subskalen ist zwar zu erwarten, dass die Werte einer Person auf den Subskalen ähnlich ausfallen, jedoch impliziert diese erhöhte Wahrscheinlichkeit im Einzelfall nicht per se die Konvergenz der Werte auf den Subskalen.

Inventar und Situational Judgement Test sind für die Messung von Veränderungen nur bedingt geeignet, da die Veränderungssensibilität bzw. Stabilität der erfassten Merkmale noch nicht empirisch überprüft wurde. Konvergenzen zwischen Selbst- und Fremdeinschätzungen auf den Skalen des Inventars liefern jedoch Hinweise auf relative Stabilität der mit dem Inventar erfassten Merkmale.

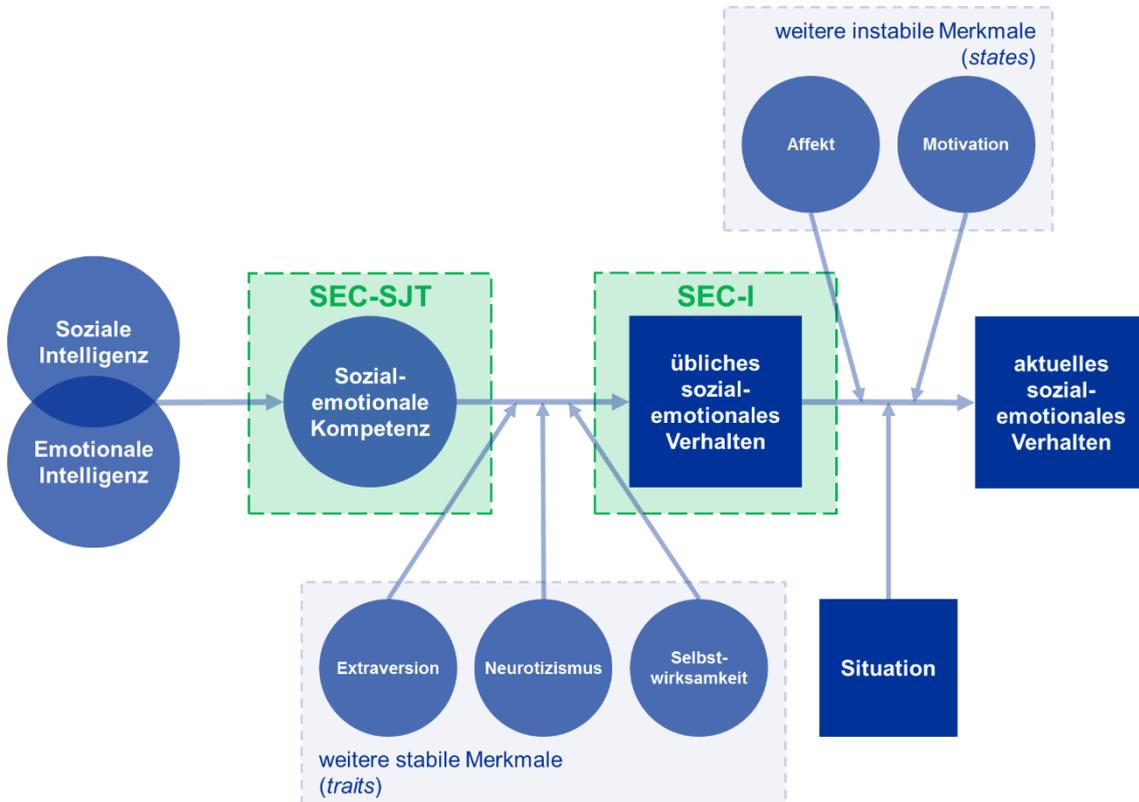
Unabhängig von den an die Erhebung anschließenden spezifischen Fördermaßnahmen kann bereits die Erhebung und die Rückmeldung von Ergebnissen als eine (weitestgehend unspezifische) Intervention angesehen werden, die der Testperson dabei helfen kann, sich selbst (besser) zu verstehen oder auch zu erkennen, wie sie von anderen wahrgenommen wird.

1.5. Theoretischer Hintergrund

Abbildung 2 (auf der nächsten Seite) veranschaulicht das zugrundeliegende konzeptuelle Verständnis des gemeinsamen Wirkens von sozial-emotionaler Kompetenz, weiteren (relativ) stabilen und instabilen Merkmalen sowie situativen Einflüssen auf aktuelles sozial-emotionales Verhalten (SEC Modell). In der Abbildung sind die von Situational Judgement Test (SEC-SJT) und Inventar (SEC-I) erfassten Merkmalsbereiche grün umrahmt.

Abbildung 2

Modell zum Einfluss sozial-emotionaler Kompetenz auf aktuelles sozial-emotionales Verhalten (SEC Modell)



Anmerkung. Bei Extraversion, Neurotizismus und Selbstwirksamkeit bzw. Affekt und Motivation handelt es sich um ausgewählte Beispiele für weitere stabile bzw. instabile Merkmale, die aktuelles sozial-emotionales Verhalten beeinflussen können.

Im SEC Modell bilden Soziale Intelligenz (SI) und Emotionale Intelligenz (EI) die Basis. EI ist die Fähigkeit, Emotionen der eigenen wie auch anderer Personen wahrzunehmen, Emotionen in das Denken einzubeziehen, Emotionen zu verstehen und zielführend mit ihnen umzugehen, und Emotionen der eigenen wie auch anderer Personen zu regulieren (Mayer et al., 2011). SI ist die Fähigkeit, andere Personen zu verstehen und mit ihnen umzugehen bzw. sich in zwischenmenschlichen Beziehungen „weise“ zu verhalten (Thorndike, 1920). EI überschneidet sich in Teilen mit SI, denn Emotionen sind in vielen sozialen Kontexten von besonderer Bedeutung (Kihlstrom & Cantor, 2011). Die Forschung zu EI hat in den 2000er Jahren die Forschung zu SI überholt, sodass SI heute mehr in der Praxis als in der Wissenschaft diskutiert wird (Kihlstrom & Cantor, 2011).

Meta-Analysen zeigen, dass EI von hoher persönlicher wie gesellschaftlicher Relevanz ist, denn sie steht in Zusammenhang mit:

- ◆ psychischer und physischer Gesundheit (Martins et al., 2010; Schutte et al., 2007),

- ◆ Zufriedenheit (Sánchez-Álvarez et al., 2016),
- ◆ dem Generalfaktor der Persönlichkeit (van der Linden et al., 2017), der als die Fähigkeit, sich sozial geschickt zu verhalten, interpretiert werden kann (van der Linden et al., 2016),
- ◆ schulischer Leistung (MacCann et al., 2019; Sánchez-Álvarez et al., 2020),
- ◆ beruflicher Leistung (Joseph et al., 2015; O’Boyle et al., 2011), wie auch
- ◆ der Einstellung zur Arbeit (Miao et al., 2017).

Intelligenz ist eine relativ breite Fähigkeit (McGrew, 2009). Im Allgemeinen sind Fähigkeiten Potentiale einer Person, also das, was eine Person unter (theoretisch) idealen Voraussetzungen leisten könnte. Prinzipiell werden Fähigkeiten (*abilities*) mit Leistungstests (*tests of maximum performance*) erfasst (Cronbach, 1990). Die Operationalisierung von EI oder SI als Fähigkeit bereitet jedoch Schwierigkeiten, sodass EI oder SI nur selten als Leistung erfasst werden (Petrides et al., 2016). Häufiger werden EI und SI mittels Fragebogen zur Selbsteinschätzung erfasst. Die Meta-Analyse von Mabe und West (1982) konnte jedoch zeigen, dass Menschen grundsätzlich nur eingeschränkt in der Lage sind, eigene Fähigkeiten objektiv einzuschätzen. Auch eine Meta-Analyse zu selbst eingeschätzter (allgemeiner) Intelligenz zeigt, dass Leistungen in Intelligenztests nur relativ geringe Zusammenhänge zur selbst eingeschätzten Intelligenz aufweisen (Freund & Kasten, 2012). Fragebögen (*tests of typical performance*) sind im Gegensatz zu Leistungstests für die Erfassung von Eigenschaften (*traits*) und Zuständen der Person (*states*) und weniger für die Erfassung von Fähigkeiten geeignet (Cronbach, 1990). Eigenschaften sind übliche Verhaltensweisen und müssen nicht zwingend dem Potential einer Person entsprechen. Wie auch bei anderen Fähigkeiten zeigt eine Meta-Analyse für EI, dass Zusammenhänge zwischen objektiv erfasster und selbst eingeschätzter EI relativ gering ausfallen (Joseph & Newman, 2010).

Meta-Analysen liefern empirische Belege dafür, dass der Umgang mit Emotionen trainiert und verbessert werden kann (Hodzic et al., 2018; Mattingly & Kraiger, 2019). Um diesen Aspekt deutlich zu machen und angesichts der Tatsache, dass relativ breite Fähigkeiten wie Intelligenz nur sehr eingeschränkt trainiert werden können, spezifische Fähigkeiten hingegen schon (Ackerman et al., 2010; Owen et al., 2010; Sala et al., 2018; Sala & Gobet, 2017), wird in diesem Testmanual dem Begriff sozial-emotionale Kompetenz gegenüber dem Begriff EI der Vorzug gegeben.

Mit dem Situational Judgement Test Paradigma eröffnen sich neue Möglichkeiten der Operationalisierung sozial-emotionaler Kompetenz als Leistung (Libbrecht & Lievens, 2012; MacCann & Roberts, 2008). Situational Judgement Tests erfassen die Leistung einer Testperson in Form von Wissen bzw. Erfahrung (im Fall sozial-emotionaler Kompetenz das

Wissen bzw. die Erfahrung, sich in verschiedensten Situationen sozial-emotional kompetent verhalten zu können) und weniger in Form einer relativ breiten Fähigkeit. Doch anders als bei (fluider) Intelligenz, die prinzipiell unabhängig von Wissen bzw. Erfahrung ist (Cattell, 1963; J. L. Horn & Cattell, 1966), profitieren spezifische Fähigkeiten durchaus von Wissen bzw. Erfahrung, wie beispielsweise die Leistungsverbesserung in geübten Intelligenztestaufgaben zeigt (Ackerman et al., 2010; Owen et al., 2010).

Im SEC Modell steht der Begriff der sozial-emotionalen Kompetenz daher für ein Bündel aus spezifischen sozial-emotionalen Fähigkeiten, die durch die Vermittlung von Wissen bzw. Erfahrung entwickelt und gefördert werden können. Um die sozial-emotionale Kompetenz multimethodal erfassen zu können, wurden speziell aufeinander abgestimmte diagnostische Verfahren für Jugendlichen und junge Erwachsene mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelt. Die sozial-emotionale Kompetenz kann mit dem Situational Judgement Test als Leistung erfasst werden. Doch alleine über das Wissen bzw. die Erfahrung zu verfügen, sich sozial-emotional kompetent verhalten zu können, muss nicht zwangsläufig im üblichen Verhalten einer Person zum Ausdruck kommen. Übliches Verhalten resultiert meist aus einem Zusammenspiel von Potential und weiteren (relativ) stabilen Merkmalen einer Person (beispielsweise Extraversion, Neurotizismus und Selbstwirksamkeit). Übliches sozial-emotionales Verhalten kann mit dem Inventar aus subjektiver Perspektive der Person (Fragebogen zur Selbsteinschätzung) wie auch aus Perspektive einer oder mehrerer anderer Personen (Fragebogen zur Fremdeinschätzung) erfasst werden. Übliches Verhalten ist im Allgemeinen ein relativ guter Anhaltspunkt für die Vorhersage aktuellen Verhaltens (Ouellette & Wood, 1998), jedoch hängt aktuelles Verhalten auch immer von situativen Einflüssen und weiteren instabilen Merkmalen der Person (beispielsweise Affekt und Motivation) ab.

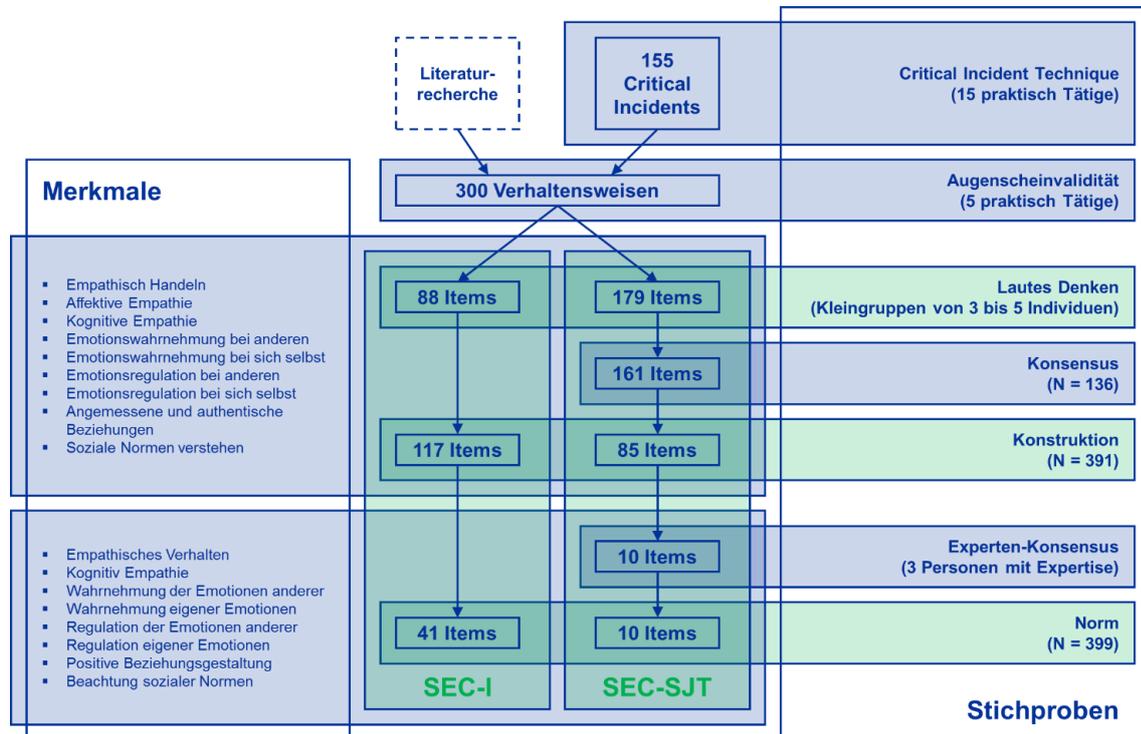
1.6. Konstruktion von Inventar und Situational Judgement Test

Die diagnostischen Verfahren wurden in einem mehrstufigen Vorgehen, in enger Abstimmung zwischen verfahrensentwickelnden Personen (nachfolgend als *Verfahrensentwickelnde* bezeichnet) und in der Betreuung bzw. Förderung von Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung praktisch tätigen Personen (nachfolgend als *praktisch Tätige* bezeichnet), entwickelt, um eine möglichst gute Anpassung an praktische Bedürfnisse und eine hohe Praktikabilität zu gewährleisten. Über die gesamte Entwicklung hinweg wurden verschiedene Konstruktionsmethoden kombiniert. Stufenweise kam eine Mischform aus Prototypenansatz, sowie deduktiver und induktiver Methode zum Einsatz. Zunächst wurde ein großer Pool an Items generiert, der anschließend mehrfach hinsichtlich der Passung für Jugendliche und junge Erwachsene

mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung untersucht, umformuliert und psychometrisch optimiert wurde. Abschließend wurden unter Berücksichtigung der Inhaltsvalidität der Skalen die psychometrisch geeignetsten Items selektiert. Abbildung 3 skizziert die Konstruktionsschritte von Inventar und Situational Judgement Test in chronologischer Reihenfolge, die im Folgenden ausführlicher beschrieben werden.

Abbildung 3

Schritte der Konstruktion des Inventars zur Erfassung sozial-emotionaler Kompetenz(en) (SEC-I) sowie des Situational Judgement Tests (SEC-SJT)



Anmerkung. Stichproben junger Menschen mit subklinischer oder klinischer kognitiver bzw. psychischer Beeinträchtigung sind grün hinterlegt.

Um möglichst viele, für sozial-emotional kompetentes Verhalten auch theoretisch relevante Verhaltensweisen zu identifizieren und verschiedene Möglichkeiten der Operationalisierung zu erschließen, wurde zunächst eine umfassende Literaturrecherche zu emotionaler und sozialer Intelligenz bzw. Kompetenz durchgeführt. Offensive Verhaltensweisen, die primär auf die Erlangung positiver Konsequenzen (ausschließlich) für die eigene Person fokussieren, wie sie beispielsweise im Inventar sozialer Kompetenz (Kanning, 2009) oder auch im Gruppentraining sozialer Kompetenz (Hinsch & Pfungsten, 2015) den sozialen Kompetenzen zugerechnet werden, sind aufgrund der geringeren prosozialen Ausrichtung nicht in die Recherche einbezogen worden.

Parallel und ergänzend zur Literaturrecherche wurden mit Hilfe einer zweckmäßig angepassten Form der Task-Analysis-Tools (TAToo) (Koch & Westhoff, 2012) relevante Merkmale sozial-emotionaler Kompetenz herausgearbeitet: Zunächst wurden nach der Methode der Critical Incident Technique (CIT) (Flanagan, 1954), detaillierte Beispiele außergewöhnlich sozial-emotional kompetenten bzw. inkompetenten Verhaltens von praktisch Tätigen schriftlich erhoben. Für jedes selbst genannte Beispiel sollten Aspekte der Situation und des Verhaltens sowie Emotionen und Gedanken detailliert beschrieben werden. Auf diese Weise konnten von 15 praktisch Tätigen aus den Beschreibungen von insgesamt 155 Critical Incidents (73 außergewöhnlich kompetenten bzw. 82 außergewöhnlich inkompetenten sozial-emotionalen Verhaltens) situative, behaviorale, emotionale und kognitive Aspekte sozial-emotional kompetenten Verhaltens abgeleitet werden. Auf Basis der Ergebnisse der Literaturrecherche und der CIT konnte ein Pool von 300 Verhaltensweisen erstellt werden. Diese wurden in einem anschließenden Workshop mit praktisch Tätigen in neun Merkmale sozial-emotionaler Kompetenz strukturiert (Empathisch Handeln, Affektive Empathie, Kognitive Empathie, Emotionswahrnehmung bei andern, Emotionswahrnehmung bei sich selbst, Emotionsregulation bei andern, Emotionsregulation bei sich selbst, Angemessene und authentische Beziehungen und Soziale Normen verstehen). Daran anschließend wurde von weiteren fünf praktisch Tätigen eine Einschätzung der Augenscheinvalidität und Angemessenheit der 300 Verhaltensweisen für die neun übergeordneten Merkmale erfasst, um wenig augenscheinvalide oder unangemessene Verhaltensweisen auszusortieren.

Auf Basis der herausgearbeiteten relevanten Merkmale sozial-emotionaler Kompetenz und des reduzierten Pools an Verhaltensweisen wurden erstmalig und weitestgehend in leichter Sprache (*Leichte Sprache: Ein Ratgeber*, 2014) 88 Items für einen Fragebogen zur Selbsteinschätzung (SR-Itempool) sowie weitere 179 Items (mit jeweils vier Antwortalternativen) für einen Situational Judgement Test (SJT-Itempool) zur Erfassung der neun Merkmale formuliert.

Um eine möglichst gute Passung der Items für Jugendliche und junge Erwachsene mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung zu erreichen, wurden die 88 Items des SR-Itempools bzw. die 179 Items des SJT-Itempools in mehreren Kleingruppen von drei bis fünf Personen des intendierten Anwendungsbereichs (insgesamt ca. 15 Personen) nach der Methode des lauten Denkens (Ericsson & Simon, 1980; Payne, 1994) bearbeitet. Die erhobenen Daten wurden den Verfahrensentwickelnden vollständig anonymisiert zur Verfügung gestellt. Auf Basis der Ergebnisse aus der *Lautes Denken Stichprobe* wurde ein erheblicher Teil der Items und auch der Antwortalternativen umformuliert, um ihre Passung für Personen des intendierten Anwendungsbereichs zu verbessern. Im Zuge dieser Überarbeitungen wurde der SR-Itempool um weitere 29 Items auf insgesamt 117 Items erhöht, der SJT-Itempool um 18 Items auf 161 Items reduziert.

Da die Bearbeitung des gesamten SJT-Itempools (161 Items mit vier Antwortalternativen) nach Einschätzung praktisch Tätiger viele Personen des intendierten Anwendungsbereichs überfordert hätte, wurden die Items zunächst in einer Stichprobe Studierender ($N = 136$) in einer vollständig anonymen Onlinestudie erhoben, um psychometrisch untersucht und angemessen reduziert werden zu können. Als Anreiz für die Teilnahme wurde unter allen Teilnehmenden ein SAMSUNG Galaxy Tab A verlost. Anhand der Daten wurden zunächst für jedes Item die modale Antwortalternative und die relativen Häufigkeiten der Antwortalternativen nach der Konsensusmethode (Barchard & Russell, 2006; Brackett & Salovey, 2006; Mayer et al., 1990; Mayer & Geher, 1996; Steinmayr et al., 2011) ermittelt. Für einen Vergleich der beiden verschiedenen Auswertungsmethoden wurden die Daten basierend auf der modalen Antwortalternative sowie den relativen Häufigkeiten kodiert, anschließend die psychometrische Güte der Items geprüft und deren faktorielle Struktur mittels Hauptkomponentenanalysen untersucht. Auf Basis der Ergebnisse aus der *Konsensusstichprobe* wurde der SJT-Itempool um 76 Items auf 85 Items reduziert.

Für den nächsten Schritt der empirischen Untersuchung wurde zunächst für jedes der 117 Items des SR-Itempools ein paralleles Item für eine Fremdeinschätzung (OR-Itempool) formuliert (geschlechtsneutrale Formulierung in der dritten, anstelle der ersten Person Singular). Der SR-Itempool und der SJT-Itempool wurden dann einer größeren *Konstruktionsstichprobe* von Personen des intendierten Anwendungsbereichs zur Bearbeitung unter späteren Erhebungsbedingungen vorgelegt ($N = 391$). Der OR-Itempool wurde jeweils einer weiteren Person (beispielsweise Auszubildende, Betreuende, Lehrkräfte), die seit mindestens einem Monat regelmäßigen Kontakt zur Person aus der Konstruktionsstichprobe hatte, zur Bearbeitung vorgelegt. Ergänzend wurden umfangreiche demografische Daten zur Person eingeholt (Alter, Geschlecht, Schulabschluss, Ausbildungsstatus, Informationen zu Art und Schwere der Beeinträchtigung, Ergebnisse bereits erfolgter psychologischer Diagnostik, berufliche und schulische Leistungen). Sämtliche Daten wurden den Verfahrensentwickelnden vollständig anonymisiert zur Auswertung übermittelt. Aufgrund fehlender Daten (teilweise unbearbeitete Items in einem Verfahren bzw. mindestens ein vollständig unbearbeitetes Verfahren) konnten vollständige Analysen über SR-Itempool, OR-Itempool und SJT-Itempool hinweg nur für 98 Datensätze gerechnet werden. Die vollständigen Analysen konnten durch weitere Teilanalysen in Substichproben ergänzt werden. Im Rahmen der Analysen wurde für jedes der drei diagnostischen Verfahren die psychometrische Güte der Items sowie deren faktorielle Struktur mittels Hauptkomponentenanalysen und Parallelanalyse nach Horn (1965) untersucht. Zusätzlich wurden die Verfahren auf konvergente, divergente und konkurrente Validität geprüft. Bei den Analysen haben sich Items zur Selbsteinschätzung, die entgegengesetzt der Merkmalsausprägung formuliert sind, trotz der Formulierung in leichter Sprache, durchgehend als problematisch erwiesen. Die Items zur Selbsteinschätzung des Merkmals Affektive Empathie haben sich nahezu vollständig in andere Merkmale eingeordnet, sodass dieses Merkmal aufgegeben wurde. Für die Items des

Situational Judgement Tests ist keine differenzierte faktorielle Struktur hervorgetreten. Auf Basis der Ergebnisse aus der *Konstruktionsstichprobe* wurde der SR-Itempool und der OR-Itempool gleichermaßen um 76 Items auf 41 parallele Items reduziert, der SJT Itempool um 75 Items auf 10 Items.

Da selbst die 10 psychometrisch geeignetsten Items des SJT-Itempools, entgegen der Einschätzung von praktisch Tätigen, von einem Großteil der Stichprobe fast vollständig gelöst werden konnten, wurde entschieden, die im SJT-Itempool verbliebenen 10 Items noch einmal zu überarbeiten und die vier Antwortalternativen um einen weiteren Distraktor zu ergänzen. Die überarbeiteten Items wurden drei verschiedenen Personen (hierbei handelte es sich nicht um die Personen, die dieses Testmanual verfasst haben) mit abgeschlossener psychologischer Ausbildung, die gegenwärtig in der Forschung tätig und mit der Fachliteratur bzw. gängigen Theorien zu sozialen und emotionalen Kompetenzen vertraut sind, zur Ermittlung von Antworten nach Expertenmethode (Barchard & Russell, 2006; Brackett & Salovey, 2006; Mayer et al., 1990; Mayer & Geher, 1996; Steinmayr et al., 2011) bzw. zur Prüfung der Übereinstimmung mit den Antworten nach Konsensumethode vorgelegt. Von den drei Personen wurde unabhängig voneinander übereinstimmend für jedes Item die gleiche Antwortalternative gewählt und die Wahl theoretisch fundiert begründet. Die Antworten nach Expertenmethode stimmen mit sämtlichen Antworten nach Konsensumethode überein.

Die im SR-Itempool, OR-Itempool und SJT-Itempool verbliebenen Items wurden abschließend als finale Versionen einer *Normstichprobe* von $N = 399$ Individuen des intendierten Anwendungsbereichs zur Bearbeitung unter späteren Erhebungsbedingungen vorgelegt. Sämtliche Ergebnisse der empirisch psychometrischen Überprüfung der finalen Versionen der diagnostischen Verfahren finden sich im empirischen Teil (Kapitel *Testtheoretische Fundierung*, ab S. 89) dieses Testmanuals.

1.7. Workshop zur Diagnostik sozial-emotionaler Kompetenzen

Im Rahmen des von der Europäischen Kommission geförderten Erasmus+-Projekts SEC4VET (Assessing and Promoting Social-Emotional Competences of Young People with Cognitive Impairments in Vocational Education and Training, Förderkennzeichen 2017-1-DE02-KA202-004136) wurde neben den diagnostischen Verfahren unter anderem auch eine Train-the-Trainer-Workshopreihe zur Fortbildung zum Themenbereich sozial-emotionaler Kompetenzen und zum Umgang mit dem Fragebogen zur Selbsteinschätzung, dem Fragebogen zur Fremdeinschätzung und dem Situational Judgement Test entwickelt (weitere Informationen hierzu finden Sie direkt auf der Homepage des Projekts unter

<https://sec4vet.eu/>). Um variabel auf den Wissenstand und die Bedürfnisse von Teilnehmenden eingehen zu können, ist der Workshop zur Diagnostik sozial-emotionaler Kompetenzen in Themenbereiche untergliedert (Schneider et al., 2020). Es werden Schulungsunterlagen und Zusatzmaterialien zu den verschiedenen Themenbereichen kostenfrei zur Verfügung gestellt. Unter anderem werden folgende Themenbereiche behandelt:

- ◆ Theoretische Grundlagen und verschiedene Definitionen sozial-emotionaler Kompetenz
- ◆ Grundlagen psychologischer Diagnostik
- ◆ Potenzielle Schwierigkeiten beim Einsatz der diagnostischen Verfahren
- ◆ Anwendung des Fragebogens zur Selbsteinschätzung bzw. des Situational Judgment Tests
- ◆ Anwendung des Fragebogens zur Fremdeinschätzung und objektive Verhaltensbeobachtung

Die verschiedenen Themenbereiche können je nach Bedarf so zusammengestellt werden, dass daraus ein individualisierter Workshop hervorgeht. Nach Einarbeitung in die jeweiligen Unterlagen zu einem der Themenbereiche sowie die betreffenden Inhalte dieses Testmanuals, kann der Workshop von Fachpersonal mit psychologischer Ausbildung durchgeführt werden, um psychologische Hilfskräfte in Durchführung und Auswertung der diagnostischen Verfahren zu schulen.

2. Anwendung

Das zweite Kapitel widmet sich der Durchführung, Auswertung und Interpretation der diagnostischen Verfahren. Der Schwerpunkt der Beschreibung der Durchführung liegt vor allem auf der standardisierten Erhebung des Fragebogens zur Selbsteinschätzung und des Situational Judgement Tests bei Personen des Anwendungsbereichs. Die Auswertung der drei diagnostischen Verfahren wird ausführlich beschrieben. Neben spezifischen Hinweisen zur Interpretation der diagnostischen Verfahren wird auch auf allgemeine Grundlagen der Interpretation psychologischer Testergebnisse eingegangen.

2.1. Anwendungsbereich

Das Inventar und der Situational Judgement Test sind speziell für die Erhebung sozial-emotionaler Kompetenz(en) bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren konstruiert und normiert. Der Fragebogen zur Selbsteinschätzung und der Situational Judgement Test sind für die Erhebung als Einzel- und als Gruppentest geeignet. Für den Fragebogen zur Selbsteinschätzung und den Fragebogen zur Fremdeinschätzung liegen *T*-Werte und Prozentränge vor, für den Situational Judgement Test Prozentränge. Die Normwerte liegen sowohl für eine Gesamtstichprobe von Jugendlichen und jungen Erwachsenen mit (sub-)klinischer Intelligenzstörung bzw. umschriebener Entwicklungsstörung als auch für Substichproben weiblichen und männlichen Geschlechts vor. Es liegt keine bevölkerungsrepräsentative Norm vor. Speziell der Situational Judgement Test kann auch als Screeningverfahren für die kriteriumsorientierte Diagnostik eingesetzt werden (siehe Abschnitt *Situational Judgement Test als Screeningverfahren*, S. 73).

Um Verständnisschwierigkeiten vorzubeugen sind die drei diagnostischen Verfahren vollständig in leichter Sprache verfasst. Alle Items des Inventars sind in Richtung der Merkmalsausprägungen formuliert. Daher wird insbesondere von einer Anwendung des Fragebogens zur Selbsteinschätzung in Kontexten, in denen ein Anlass zur Verfälschung bestehen könnte, abgeraten. Empirische Befunde bestätigen, dass die Ergebnisse aus dem Fragebogen zur Selbsteinschätzung und dem Situational Judgement Test nicht vom Vorliegen einer körperlichen Behinderung (siehe Abschnitt *Divergente Validität*, S. 106) sowie die Ergebnisse aus dem Fragebogen zur Selbsteinschätzung und dem Fragebogen zur Fremdeinschätzung nicht vom Vorliegen einer geistigen Behinderung (siehe Abschnitt *Extremgruppenvalidierung*, S. 108) beeinträchtigt werden.

Die Erhebung und Auswertung des Fragebogens zur Selbsteinschätzung und des Situational Judgement Tests sind relativ einfach und können von geschulten psychologischen Hilfskräften übernommen werden. Der Fragebogen zur Selbsteinschätzung sollte nur von Personen ausgefüllt werden, die mit der einzuschätzenden Person seit mindestens einem Monat regelmäßigen Kontakt haben (beispielsweise Auszubildende, Betreuende, Lehrkräfte),

sodass sich bereits ausreichende Gelegenheiten ergeben haben, übliches Verhalten der Testperson in verschiedenen Situationen zu beobachten. Empirische Befunde bestätigen, dass die Ergebnisse aus dem Fragebogen zur Fremdeinschätzung nicht von der Dauer der Bekanntheit beeinträchtigt werden (siehe Abschnitt *Divergente Validität*, S. 106).

Das Formulieren einer angemessenen diagnostischen Frage, die Wahl einer Norm, die Wahl einer Vertrauenswahrscheinlichkeit oder das Abwägen von Nutzen und Risiken verschiedener Entscheidungen erfordern fundierte Fachkenntnisse und einschlägige Erfahrung in der psychologischen Diagnostik. Da alldem kein allgemeingültiges Schema zugrunde liegt, können an manchen Stellen auch keine einfachen Empfehlungen gegeben werden. Die Interpretation sollte daher Fachpersonal mit psychologischer Ausbildung vorbehalten bleiben (siehe Abschnitt *Qualifikation der interpretierenden Person*, S. 50).

2.2. Durchführung

Die Inhalte dieses Abschnitts befassen sich primär mit der Erhebung des Fragebogens zur Selbsteinschätzung bzw. des Situational Judgement Tests bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung. Besonderheiten des Fragebogens zur Fremdeinschätzung und objektiver Verhaltensbeobachtung widmet sich gesondert einer der Themenbereiche aus dem *Workshop zur Diagnostik sozial-emotionaler Kompetenzen* (siehe S. 25).

2.2.1. Qualifikation der Testleitung

Die Erhebung ist soweit standardisiert, dass sie von geschulten psychologischen Hilfskräften vorgenommen werden kann, die mit den hierfür relevanten Inhalten dieses Testmanuals vertraut sind. Insbesondere der Abschnitt zur *Durchführung* sollte gelesen worden sein. Darüber hinaus sollte sich die Testleitung anhand sämtlicher Testmaterialien in die praktische Handhabung eingearbeitet haben. Sowohl die Kenntnis gängiger internationaler Standards für psychologisches Testen (Häcker et al., 1998; International Test Commission, 2001) wie auch Vorerfahrungen in der Durchführung psychologisch diagnostischer Erhebungen sind hierbei von Vorteil.

Zusätzlich können psychologische Hilfskräfte mit dem *Workshop zur Diagnostik sozial-emotionaler Kompetenzen* (siehe S. 25) in der Durchführung und Auswertung der diagnostischen Verfahren geschult werden.

2.2.2. Anforderungen an die Testperson

Die wesentlichste Anforderung an die Testperson ist eine positive Einstellung gegenüber der Erhebung. Diese betrifft sowohl die Motivation zur freiwilligen Teilnahme wie auch die Bereitschaft ehrlich zu antworten. Ausdrücklich sei darauf hingewiesen, dass die Erhebung einer Testperson ausschließlich auf freiwilliger Basis erfolgen sollte. Der Versuch, eine Testperson zur Erhebung zu überreden oder gar zu zwingen, ist nicht zielführend und sollte unterbleiben. Die Bearbeitung des Fragebogens zur Selbsteinschätzung erfordert von der Testperson zusätzlich eine gewisse Fähigkeit zur Selbstreflexion, die des Situational Judgement Tests die Fähigkeit über einen zeitlich begrenzten Zeitraum konzentriert arbeiten zu können. Zum gegenwärtigen Zeitpunkt liegen keine Hinweise vor, die auf eine Beeinträchtigung der Ergebnisse durch das Vorliegen einer körperlichen Behinderung hindeuten (siehe Abschnitt *Divergente Validität*, S. 106). Bei körperlichen Beeinträchtigungen wie auch bei Beeinträchtigungen des Sehvermögens können die Items vorgelesen und die Antworten der Testperson übernommen werden.

2.2.3. Allgemeine Hinweise zur Erhebung

Die Bearbeitung des Fragebogens zur Selbsteinschätzung und des Situational Judgement Tests erfolgt ohne Zeitlimit. Erfahrungsgemäß nimmt die Bearbeitung des Inventars (Fragebogen zur Selbsteinschätzung oder Fragebogen zur Fremdeinschätzung) in etwa 10 Minuten, die des Situational Judgement Tests in etwa 30 Minuten bzw. in Einzelfällen auch bis zu 60 Minuten in Anspruch. Die Antworten auf die Items werden direkt in die Testmaterialien eingetragen. Bei körperlichen Beeinträchtigungen wie auch bei Beeinträchtigungen des Sehvermögens ist entsprechende Unterstützung möglich, beispielsweise das Vorlesen der Items sowie Hilfe beim Ausfüllen. Wird Unterstützung gegeben, sollte sich diese ebenfalls streng an der standardisierten Durchführung (siehe Abschnitt *Gestaltung der Erhebungssituation*, S. 31; Abschnitt *Instruktion der Testperson(en)*, S. 32; und Abschnitt *Umgang mit Fragen und Störungen während der Erhebung*, S. 33) und den Testmaterialien orientieren, sich also auf Hilfe bei der Bearbeitung beschränken und beispielsweise nicht bei der Auswahl einer passenden Antwort unterstützen. Von Umformulierungen der Instruktion oder auch einzelner Items sollte abgesehen werden.

Der Fragebogen zur Selbsteinschätzung und der Situational Judgement Test können als Einzel- und als Gruppentest durchgeführt werden. Bei Gruppentests sollte die Gruppengröße überschaubar gehalten werden, um gegenseitigen Störungen vorzubeugen. Testpersonen, die bereits früher Auffälligkeiten im Verhalten gezeigt haben, sodass von ihnen Störungen des Gruppentests zu erwarten sind, wie auch Personen, die aufgrund einer Beeinträchtigung Unterstützung benötigen, sollten stets einzeln getestet werden. Bei Gruppentests sollte darauf geachtet werden, dass Testmaterialien nicht durcheinanderkommen

und zu jeder Zeit den einzelnen Testpersonen zugeordnet werden können. Es ist daher ratsam, dass die Angaben zur Testperson (Name und Vorname oder Personenkennziffer, Alter und Geschlecht) immer auf der ersten Seite der entsprechenden Testmaterialien eingetragen werden, idealerweise bereits vorab durch die Testleitung. Dabei sollte selbstverständlich sein, dass psychologische Testergebnisse vertraulich zu behandeln und nicht weiterzugeben sind. Trägt die Testleitung die Angaben zur Testperson bereits vorab ein, kann die Testperson gebeten werden, die Angaben zu prüfen. Dadurch kann bei Gruppentests sichergestellt werden, dass jede Testperson die für sie vorgesehenen Unterlagen erhalten hat.

Auf der ersten Seite des Fragebogens zur Fremdeinschätzung können zusätzliche Angaben zur Person, die die Fremdeinschätzung abgibt, eingetragen werden (Name und Vorname, Tätigkeitsbereich und Dauer der Bekanntheit). Fremdeinschätzungen sollten nur dann abgegeben werden, wenn die Person, die die Einschätzung abgibt, bereits seit längerer Zeit regelmäßigen Kontakt zur einzuschätzenden Person hat (mindestens seit einem Monat) und ausreichend Gelegenheit hatte, übliches Verhalten der Testperson in verschiedenen Situationen zu beobachten.

Die folgenden Abschnitte zur *Gestaltung der Erhebungssituation* (siehe S. 31), der *Instruktion der Testperson(en)* (siehe S. 32) und zum *Umgang mit Fragen und Störungen während der Erhebung* (siehe S. 33) sollten von der Testleitung soweit verinnerlicht werden, dass die Standardisierung der Durchführung sichergestellt werden kann.

2.2.4. Gestaltung der Erhebungssituation

Die Testleitung sollte sich darum bemühen, eine möglichst angenehme Erhebungssituation zu schaffen, die sich förderlich auf eine positive Einstellung der Testperson gegenüber der Erhebung auswirken kann. Für den Fragebogen zur Selbsteinschätzung gilt, dass die Testpersonen dazu motiviert werden sollten, möglichst offen und ehrlich zu antworten. Eine Prüfungsatmosphäre sollte unbedingt vermieden werden, da hierdurch unter Umständen ein Motiv zur Täuschung gefördert werden oder die Erhebung abgelehnt werden könnte.

Folgende Maßnahmen fördern eine angenehme Erhebungssituation für die Testperson und sollten entsprechend umgesetzt werden:

- ◆ Der Testperson sollte Sinn und Zweck der Erhebung erläutert werden.
- ◆ Der Testperson sollte zugesichert werden, dass die Ergebnisse der Erhebung streng vertraulich sind.
- ◆ Der Testperson sollte ein ausführliches Anschlussgespräch mit Rückmeldung angeboten werden, sofern sie sich ein solches Gespräch wünscht.

- ◆ Die Testleitung sollte sich ausreichend Zeit für die Erhebung und die individuellen Bedürfnisse der Testperson nehmen können.
- ◆ Der Testperson sollte so viel Bearbeitungszeit wie benötigt zur Verfügung gestellt werden. Die Bearbeitung erfolgt prinzipiell ohne Zeitlimit.
- ◆ Die Instruktion sollte gemeinsam mit der Testperson durchgegangen werden und es sollte sichergestellt werden, dass alles verstanden wurde.
- ◆ Die Testperson sollte darauf hingewiesen werden, dass es keine richtigen und falschen Antworten gibt, sondern dass es darum geht, sich selbst so gut wie möglich einzuschätzen.
- ◆ Die Testperson sollte während der Erhebung nicht allein gelassen werden, allerdings auch nicht den Eindruck haben, unter strenger Beobachtung zu stehen.
- ◆ Während der Erhebung sollten Störungen unbedingt vermieden werden. Treten unvorhergesehen Störungen (beispielsweise Lärm oder Unterbrechungen) auf, sollten diese möglichst schnell beseitigt werden, sodass die Erhebung ungestört fortgesetzt werden kann.

Eine detaillierte Beschreibung der Instruktion der Testperson findet sich im anschließenden Abschnitt (siehe Abschnitt *Instruktion der Testperson(en)*, S. 32). Falls die Testperson während der Erhebung Fragen hat oder es zu Störungen kommt, finden sich Hinweise zum Umgang mit diesen im Abschnitt *Umgang mit Fragen und Störungen während der Erhebung* (siehe S. 33).

Konnten während der Erhebung Auffälligkeiten im Verhalten der Testperson beobachtet werden, sollten diese von der Testleitung in freier Form notiert und nach Beendigung der Bearbeitung in den Testmaterialien schriftlich dokumentiert werden (hierfür ist der freie Bereich unter der Überschrift *Beobachtungen der Testleitung* auf der jeweils letzten Seite der entsprechenden Testmaterialien vorgesehen). Auffälligkeiten liegen vor, wenn eine Testperson beispielsweise mit der Bearbeitung aller oder eines Teils der Items schnell fertig ist, während der Bearbeitung viel nachfragt oder unruhig/unkonzentriert wirkt.

2.2.5. Instruktion der Testperson(en)

Die standardisierte Instruktion für den Fragebogen zur Selbsteinschätzung bzw. den Situational Judgement Test sowie jeweils ein Beispielitem sind auf den ersten beiden Seiten der entsprechenden Testmaterialien abgedruckt. Die gesamte Instruktion sollte der Testperson laut vorgelesen werden. Die Testperson kann dabei selbständig mitlesen. Das Beispielitem sollte gemeinsam mit der Testperson durchgegangen werden. Während des Vorlesens sollte nach jeder der im Text enthaltenen Fragen kurz innegehalten werden, um der Testperson das Antworten zu ermöglichen. Währenddessen sollte Blickkontakt aufgenommen

werden, sodass die Testperson sich angesprochen fühlt und beispielsweise auch ein Kopfschütteln der Testperson bemerkt werden kann. Die gewählten Antworten werden angekreuzt, falsche Kreuze sollten deutlich korrigiert werden, sodass eindeutig klar ist, welche Antwort gewählt wurde (dies kann am Beispielitem geübt werden).

Unmittelbar nach Beendigung der Bearbeitung des Fragebogens zur Selbsteinschätzung oder des Situational Judgement Tests sollte geprüft werden, ob alle Items von der Testperson beantwortet wurden. Sind einzelne Items unbeantwortet geblieben, sollte die Testperson noch einmal darum gebeten werden, den Versuch zu unternehmen, die fehlenden Antworten zu ergänzen. Kommen dabei Fragen auf, sollte möglichst ebenso wie bei Fragen während der Erhebung reagiert werden (siehe Abschnitt *Umgang mit Fragen und Störungen während der Erhebung*, S. 33).

Bei der Instruktion für den Fragebogen zur Selbsteinschätzung sollten einige Besonderheiten beachtet werden: Testpersonen sollten bereits ganz zu Beginn darauf hingewiesen werden, dass der Fragebogen zur Selbsteinschätzung Verhalten und Gedanken in sozialen und emotionalen Situationen erhebt und sich daher einzelne Aussagen und Fragen ähnlich sein können. Testpersonen gegenüber sollte besser von Aussagen und Fragen anstatt von Items gesprochen werden. Den Testpersonen sollte erklärt werden, dass niemand in den Kopf anderer Menschen blicken kann. Deshalb kann der Fragebogen dabei helfen, zu verstehen, was in anderen Menschen vor sich geht. Das kann der Fragebogen jedoch nur, wenn offen und ehrlich geantwortet wird. Werden die Aussagen und Fragen nicht ehrlich beantwortet, ist die Durchführung auch nicht zielführend.

Mit der Bearbeitung des Fragebogens zur Selbsteinschätzung sollte erst bzw. nur begonnen werden, wenn die Testperson bereit ist, offen und ehrlich zu antworten und keine Fragen mehr hat. Stellt die Testperson während der Erhebung Fragen, sollten diese möglichst so, wie im folgenden Abschnitt beschrieben, beantwortet werden (siehe Abschnitt *Umgang mit Fragen und Störungen während der Erhebung*, S 33).

2.2.6. Umgang mit Fragen und Störungen während der Erhebung

Kommen während der Bearbeitung des Fragebogens zur Selbsteinschätzung bzw. des Situational Judgement Tests inhaltliche Fragen auf, sollte von Umformulierungen oder Erklärungen einzelner Items abgesehen werden, sodass äußere unstandardisierte Einflüsse auf die Beantwortung der betroffenen Items minimiert werden.

Bei Fragen im Verlauf der Bearbeitung des Fragebogens zur Selbsteinschätzung kann der Testperson geantwortet werden, dass die Aussagen und Fragen absichtlich allgemein gehalten wurden, um von möglichst vielen verschiedenen Testpersonen beantwortet werden

zu können. Von Umformulierungen oder Erklärungen einzelner Items sollte abgesehen werden. Es kann auch noch einmal darauf hingewiesen werden, dass es keine richtigen und falschen Antworten gibt, sondern dass es darum geht, sich selbst so gut wie möglich einzuschätzen. Fällt es einer Testperson offensichtlich schwer, sich bei einem Item so gut wie möglich einzuschätzen, kann sie die für sie noch am ehesten passende Antwort wählen.

Kommen während der Bearbeitung des Situational Judgement Test Fragen auf, sollte die Testperson daran erinnert werden, dass es darum geht, die für sie beste bzw. noch am besten passende Antwort zu wählen. Auch hier sollte von Umformulierungen oder Erklärungen einzelner Items abgesehen werden. Bei Gruppentests sollten Fragen im direkten Kontakt mit der fragestellenden Testperson beantwortet werden, sodass andere Testpersonen möglichst nicht gestört werden.

Auftretende Störungen während der Erhebung oder auch vollständige Unterbrechungen der Erhebung sollten vermieden werden. Tritt eine unvermeidbarere Störung oder Unterbrechung ein, sollten derartige Vorkommnisse genau wie Auffälligkeiten im Verhalten der Testperson nach Beendigung der Bearbeitung in den Testmaterialien schriftlich dokumentiert werden (hierfür ist der freie Bereich unter der Überschrift *Beobachtungen der Testleitung* auf der jeweils letzten Seite der Testmaterialien vorgesehen).

2.3. Auswertung

2.3.1. Qualifikation der auswertenden Person

Genau wie die Erhebung ist die Auswertung der drei diagnostischen Verfahren soweit standardisiert, dass sie von geschulten psychologischen Hilfskräften vorgenommen werden kann. Die Hilfskräfte sollten mit den hierfür relevanten Inhalten dieses Testmanuals vertraut sein. Bei der Auswertung sollte besonders sorgfältig gearbeitet werden. Insbesondere die Auswertung von Hand ist, trotz der Standardisierung, bei unsorgfältiger Arbeitsweise anfällig für Übertragungs- oder Rechenfehler. Erfolgt die Auswertung computerbasiert (siehe Abschnitt *Computerbasierte Auswertung*, S. 50), lässt sich die Anfälligkeit für Fehler deutlich reduzieren. Die Antworten der Testperson müssen dann lediglich in die Auswertungsdatei übertragen werden, die Auswertung erfolgt vollständig automatisiert.

2.3.2. Allgemeine Hinweise zur Auswertung

Die Auswertung kann von Hand (siehe Abschnitt *Auswertung von Hand*, S. 40) oder computerbasiert (siehe Abschnitt *Computerbasierte Auswertung*, S. 50) vorgenommen werden. Die Auswertung von Hand erfolgt mit Hilfe des jeweiligen Auswertungsbogens (für den

Fragebogen zur Selbsteinschätzung mit dem *SEC-I-SR Auswertungsbogen*, für den Fragebogen zur Fremdeinschätzung mit dem *SEC-I-OR Auswertungsbogen* und für den Situational Judgement Test mit dem *SEC-SJT Auswertungsbogen*) und des Profilbogens, in dem die Ergebnisse aus dem Fragebogen zur Selbsteinschätzung, dem Fragebogen zur Fremdeinschätzung und dem Situational Judgement Test gemeinsam dokumentiert werden. Für die computerbasierte Auswertung wird eine Auswertungsdatei (Microsoft® Excel® Arbeitsmappe) zur Verfügung gestellt, in die die Antworten der Testperson der Reihenfolge nach übertragen werden.

Für Personen, die bereits Routine in der Auswertung psychologischer Tests haben, nimmt die Auswertung von Hand erfahrungsgemäß nicht mehr als 15 Minuten in Anspruch, die computerbasierte Auswertung nicht mehr als 5 Minuten. Die Auswertung von Hand wird in diesem Testmanual ausführlich erläutert. Sowohl auf dem jeweiligen Auswertungsbogen wie auch auf dem Profilbogen werden sämtliche erforderlichen Schritte (zur Auswertung, Ermittlung der Normwerte und Darstellung des Profils) noch einmal in kurzen Stichpunkten genannt. Das alleinige Lesen dieser kurzen Stichpunkte ist jedoch kein Ersatz für die ausführliche Einarbeitung in die Auswertung anhand der relevanten Abschnitte dieses Testmanuals. Die kurzen Stichpunkte behandeln beispielsweise auch nicht die zunächst erforderliche *Prüfung der Qualität der Bearbeitung* (siehe S. 35) oder die *Interpolation des Rohwerts der Globalskala des Fragebogens zur Selbsteinschätzung* oder der Globalskala des Fragebogens zur Fremdeinschätzung (siehe S. 42).

Da die computerbasierte Auswertung deutlich weniger Zeit in Anspruch nimmt und auch weniger anfällig für Fehler ist, sollte diese der Auswertung von Hand vorgezogen werden. Die computerbasierte Auswertung erfolgt analog zur Auswertung von Hand. Die automatisierten Schritte der computerbasierten Auswertung können anhand der ausführlichen Erläuterungen der Auswertung von Hand nachvollzogen werden (siehe Abschnitt *Auswertung von Hand*, S. 40).

Die rein rechnerische Auswertung, ob von Hand oder computerbasiert, setzt zunächst eine Prüfung der Qualität der Bearbeitung voraus (siehe Abschnitt *Prüfung der Qualität der Bearbeitung*, S. 35) und darf auf keinen Fall mit der Interpretation des Ergebnisses (siehe Abschnitt *Interpretation*, S. 50) verwechselt werden. Denn Rohwerte können, im Gegensatz zu *T*-Werten und Prozenträngen, nicht sinnvoll interpretiert werden (siehe Abschnitt *Standardisierte Werte*, S. 53).

2.3.3. Prüfung der Qualität der Bearbeitung

Testergebnisse aus Fragebögen zur Selbsteinschätzung sind, anders als Testergebnisse aus Leistungstest, besonders anfällig für verschiedene andere Einflüsse. Nicht mit den zu erfassen-

senden Merkmalen in Zusammenhang stehende Einflüsse können Testergebnisse aus Fragebögen (positiv) verzerren (siehe Abschnitt *Verzerrungen im Antwortverhalten*, S. 36). Bevor mit der rechnerischen Auswertung begonnen wird, ist es daher ratsam, insbesondere die Qualität der Bearbeitung des Fragebogens zur Selbsteinschätzung zu prüfen. Aber auch für die Auswertung des Fragebogens zur Fremdeinschätzung und des Situational Judgment Tests gilt, dass zumindest auf Auffälligkeiten geachtet werden sollte.

2.3.3.1. Auffälligkeiten im Verhalten der Testperson während der Erhebung

Wenn Auffälligkeiten im Verhalten der Testperson während der Erhebung beobachtet werden konnten, sollten diese nach Beendigung der Bearbeitung in den Testmaterialien schriftlich dokumentiert werden (siehe Abschnitt *Gestaltung der Erhebungssituation*, S. 31; und Abschnitt *Umgang mit Fragen und Störungen während der Erhebung*, S. 33).

Vor Beginn der Auswertung sollte zunächst geprüft werden, ob in den Testmaterialien auf der jeweils letzten Seite Beobachtungen der Testleitung eingetragen wurden. Ist es zu Auffälligkeiten während der Erhebung gekommen oder zeigen die bearbeiteten Testmaterialien anderweitige Auffälligkeiten (wenn beispielsweise Kästchen mit zunehmendem Fortschritt unsorgfältiger angekreuzt oder freie Bereiche für Kritzeleien verwendet wurden) sollte die Auswertung sorgfältig abgewogen werden. In solchen Fällen ist die Bearbeitung nicht per se ungültig, jedoch sollte im Fall von eindeutigen Zweifeln an der Qualität der Bearbeitung (beispielsweise bei durchgehend extrem hohen Antworten in der Selbsteinschätzung, die deutlich von einer ausdifferenzierten Fremdeinschätzung abweichen) von der Auswertung abgesehen werden.

2.3.3.2. Verzerrungen im Antwortverhalten

Der Fragebogen zur Selbsteinschätzung ist speziell für den Gebrauch bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung konzipiert. Daher kommen nur Items zum Einsatz, die in Richtung der Merkmalsausprägungen formuliert sind. Während der Konstruktion hat sich gezeigt, dass Items, die entgegengesetzt der Merkmalsausprägung formuliert sind, trotz einer Formulierung in leichter Sprache, zu Verständnisschwierigkeiten und Irritation bei den Jugendlichen und jungen Erwachsenen führen. Aufgrund möglicher Verständnisschwierigkeiten wurde auch auf den Gebrauch von Kontrollitems zur Prüfung des Antwortverhaltens abgesehen, deren Validität in der psychologischen Diagnostik zudem stark diskutiert wird. Entsprechend besteht eine besondere Anfälligkeit der Rohwerte für Antworttendenzen. Unter Antworttendenzen

(*response sets*) wird eine bewusste oder unbewusste systematische Verzerrung im Antwortverhalten der Testperson verstanden (Cronbach, 1946, 1950). Es können formale und inhaltliche Antworttendenzen unterschieden werden.

Systematische Verzerrungen im Antwortverhalten können für den Fragebogen zur Selbsteinschätzung aber auch für den Fragebogen zur Fremdeinschätzung per se nicht verhindert werden. Es besteht jedoch die Möglichkeit, Verzerrungen vorzubeugen, indem den Jugendlichen und jungen Erwachsenen gegenüber betont wird, wie wichtig die sorgfältige Bearbeitung sowie offene und ehrliche Antworten sind (siehe Abschnitt *Instruktion der Testperson(en)*, S. 32). Es sollte davon abgesehen werden, Testpersonen anzudrohen, dass unehrliche Antworten erkannt werden können, um sie zu ehrlichen Antworten zu bewegen. Dadurch könnten sich Testpersonen unnötig unter Druck gesetzt oder erst recht zu unehrlichen Antworten provoziert fühlen. Verzerrende Einflüsse auf Fremdeinschätzungen behandelt auch der Themenbereich *Einsatz des Fragebogens zur Fremdeinschätzung und objektive Verhaltensbeobachtung* aus dem *Workshop zur Diagnostik sozial-emotionaler Kompetenzen* (siehe S. 25).

Formale Antworttendenzen

Gängige formale Antworttendenzen (Cronbach, 1946, 1950) sind:

- ◆ Tendenz zu extremen Antworten: Personen, die zu extremen Antworten tendieren, neigen dazu, besonders häufig den höchsten oder niedrigsten Wert der Antwortskala zu wählen.
- ◆ Ja-/Nein-Sage-Tendenz: Personen, die dieses Antwortverhalten zeigen, neigen dazu, Items besonders häufig zustimmend bzw. ablehnend zu beantworten.
- ◆ Tendenz zur Mitte: Personen, die zur Mitte tendieren, neigen dazu, besonders häufig den mittleren Wert der Antwortskala zu wählen.
- ◆ Tendenz zu zufälligen Antworten: Personen, die dieses Antwortverhalten zeigen, neigen dazu, besonders häufig rein zufällig Werte der Antwortskala zu wählen.

Beispiele für sowohl die Tendenz zu extremen Antworten als auch die Tendenz zur Mitte werden in Abbildung 4 (auf der nächsten Seite) gezeigt, Beispiele für Tendenzen zu zufälligen Antworten in Abbildung 5 (auf der übernächsten Seite).

Abbildung 4

Beispiele für formale Antworttendenzen

SEC-I-SR

Bitte kreuze für jede Aussage die Abstufung an, die dich am besten beschreibt.

	nie	selten	gelegentlich	häufig	immer
1) Ich versuche zu helfen, wenn andere ängstlich sind.	X				
2) Ich kann andere beruhigen, wenn sie ängstlich sind.		X			
3) Ich bemerke, wenn ich stolz bin.	X				
4) Ich kann mich selbst beruhigen, wenn ich angespannt bin.	X				
5) Ich will mich gleich nach einem Streit wieder vertragen.	X				
6) Ich bemerke, wenn andere schlecht gelaunt sind.		X			
7) Ich kann verstehen, weshalb andere angespannt sind.	X				
8) Ich bedanke mich, wenn andere etwas für mich tun.	X				
9) Ich unterstütze andere.	X				
10) Ich versuche anderen zu helfen, wenn ich das Gefühl habe, dass sie Hilfe brauchen.	X				
11) Ich kann andere beruhigen, wenn sie angespannt sind.	X				
12) Ich bemerke, wenn ich traurig bin.	X				
13) Ich kann mich selbst beruhigen, wenn ich ängstlich bin.	X				
14) Es beschäftigt mich, wenn ich Streit mit anderen habe.	X				

Erasmus+ 4 von 7

SEC-I-SR

Bitte kreuze für jede Aussage die Abstufung an, die dich am besten beschreibt.

	nie	selten	gelegentlich	häufig	immer
1) Ich versuche zu helfen, wenn andere ängstlich sind.			X		
2) Ich kann andere beruhigen, wenn sie ängstlich sind.			X		
3) Ich bemerke, wenn ich stolz bin.			X		
4) Ich kann mich selbst beruhigen, wenn ich angespannt bin.			X		
5) Ich will mich gleich nach einem Streit wieder vertragen.			X		
6) Ich bemerke, wenn andere schlecht gelaunt sind.			X		
7) Ich kann verstehen, weshalb andere angespannt sind.			X		
8) Ich bedanke mich, wenn andere etwas für mich tun.			X		
9) Ich unterstütze andere.			X		
10) Ich versuche anderen zu helfen, wenn ich das Gefühl habe, dass sie Hilfe brauchen.			X		
11) Ich kann andere beruhigen, wenn sie angespannt sind.			X		
12) Ich bemerke, wenn ich traurig bin.			X		
13) Ich kann mich selbst beruhigen, wenn ich ängstlich bin.		X			
14) Es beschäftigt mich, wenn ich Streit mit anderen habe.			X		

Erasmus+ 4 von 7

Anmerkung. Linke Seite: Tendenz zu extremen Antworten; Rechte Seite: Tendenz zur Mitte. Bei beiden Beispielen wird jeweils nur die erste Seite aus dem bearbeiteten Fragebogen zur Selbsteinschätzung (SEC-I-SR) dargestellt.

Während Tendenzen zu extremen Antworten oder Tendenzen zur Mitte bei visueller Inspektion besonders augenfällig sein dürften (siehe Abbildung 4), ist es aufgrund des ausschließlichen Einsatzes von Items, die in Richtung der Merkmalsausprägungen formuliert sind, deutlich schwerer Ja-/Nein-Sage-Tendenzen oder Tendenzen zu zufälligem Antworten zu erkennen (siehe Abbildung 5). Tendenzen zu zufälligem Antworten sind in der Regel nur dann zu erkennen, wenn die zufälligen Antworten auf die Items einem bestimmten Muster folgen, wie beispielsweise auf der rechten Seite von Abbildung 5.

Ausgefüllte Fragebögen sollten auf jeden Fall einer visuellen Inspektion unterzogen werden, in der auf mögliche Muster in den Antworten der Testperson geachtet wird. Auch eine deutlich ausgeprägte durchgehende Divergenz zwischen Selbsteinschätzung und Fremdeinschätzung(en) kann Hinweise auf das Vorhandensein von Antworttendenzen liefern. Können solche Muster entdeckt werden, ist dies zwar kein eindeutiger Beleg für das Vorliegen einer formalen Antworttendenz, doch bei Beobachtung eines auffälligen Musters sollte die Auswertung besser nur unter Vorbehalt erfolgen.

Abbildung 5

Beispiele für formale Antworttendenzen

The image displays two versions of the SEC-I-SR questionnaire, each with 14 items and a 5-point Likert scale (nie, selten, gelegentlich, häufig, immer). The left version shows a tendency to select 'häufig' (often) for most items, while the right version shows a tendency to select 'nie' (never) for most items.

Item	nie	selten	gelegentlich	häufig	immer
1) Ich versuche zu helfen, wenn andere ängstlich sind.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2) Ich kann andere beruhigen, wenn sie ängstlich sind.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3) Ich bemerke, wenn ich stolz bin.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4) Ich kann mich selbst beruhigen, wenn ich angespannt bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
5) Ich will mich gleich nach einem Streit wieder vertragen.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6) Ich bemerke, wenn andere schlecht gelaunt sind.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
7) Ich kann verstehen, weshalb andere angespannt sind.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8) Ich bedanke mich, wenn andere etwas für mich tun.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
9) Ich unterstütze andere.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
10) Ich versuche anderen zu helfen, wenn ich das Gefühl habe, dass sie Hilfe brauchen.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11) Ich kann andere beruhigen, wenn sie angespannt sind.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
12) Ich bemerke, wenn ich traurig bin.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
13) Ich kann mich selbst beruhigen, wenn ich ängstlich bin.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14) Es beschäftigt mich, wenn ich Streit mit anderen habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

Anmerkung. Auf beiden Seiten: Tendenz zu zufälligen Antworten. Bei beiden Beispielen wird jeweils nur die erste Seite aus dem bearbeiteten Fragebogen zur Selbsteinschätzung (SEC-I-SR) dargestellt.

Inhaltliche Antworttendenzen

Neben formalen Antworttendenzen sind auch inhaltliche Verzerrungen des Antwortverhaltens im Rahmen der Erhebung möglich. Gängige inhaltliche Antworttendenzen (Paulhus, 2002) sind:

- ◆ **Soziale Erwünschtheit:** Personen, die sozial erwünscht antworten, neigen dazu, ihre Antworten besonders häufig an sozialen Standards zu orientieren, um dem gesellschaftlichen Ideal möglichst gut zu entsprechen.
- ◆ **Verfälschung:** Personen, die ihre Antworten verfälschen, neigen dazu, ihre Antworten besonders häufig am Anlass der Erhebung zu orientieren, um möglichst das zu erreichen, was sie sich von einem für sie erfolgreichen Ausgang der Erhebung erwarten.

Inhaltliche Antworttendenzen betreffen insbesondere den Fragebogen zur Selbsteinschätzung. Anders als formale Antworttendenzen, die mit einer visuellen Inspektion potentiell erkannt werden können, sind inhaltliche Antworttendenzen nicht zu erkennen. Am ehesten

kann inhaltlichen Antworttendenzen durch eine positive Einstellung der Testperson gegenüber dem Sinn und Zweck der Erhebung sowie einem persönlichen Interesse an einem ehrlichen Ergebnis vorgebeugt werden (siehe Abschnitt *Ergebnisrückmeldungen*, S. 75).

2.3.3.3. Fehlende Antworten

Am besten sollte unmittelbar nach Beendigung der Erhebung jedes der diagnostischen Verfahren geprüft werden, ob alle Items von der Testperson bzw. der einschätzenden Person beantwortet wurden. Falls dies nicht geschehen ist und Items unbeantwortet geblieben sind, besteht noch die Möglichkeit, die Testperson bzw. die einschätzende Person im Nachhinein zu bitten, fehlende Antworten zu ergänzen.

Ist die Testperson verunsichert und weiß nicht, wie sie fehlende Items beantworten kann, kann sie noch einmal daran erinnert werden, dass es keine richtigen oder falschen Antworten gibt und es darum geht, sich selbst so gut wie möglich einzuschätzen (Fragebogen zur Selbsteinschätzung) bzw. die für sie (noch) am besten passende Antwort zu wählen (Situational Judgement Test) (siehe Abschnitt *Umgang mit Fragen und Störungen während der Erhebung*, S. 33).

Fehlende Antworten sollten im Nachhinein nur dann eingeholt werden, wenn die Erhebung unmittelbar oder nur wenige Tage zurückliegt und relativ wenige Items unbeantwortet geblieben sind. Ist dies nicht der Fall, kann bei bis zu maximal vier fehlenden Antworten (insgesamt unter 10.0% fehlende Antworten) für die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) oder die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) mittels Interpolation ein Rohwert geschätzt werden (siehe Abschnitt *Interpolation des Rohwerts der Globalskala des Fragebogens zur Selbsteinschätzung oder der Globalskala des Fragebogens zur Fremdeinschätzung*, S. 42). Die Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung können, aufgrund der deutlich geringeren Anzahl an Items, ebenso wie die Globalskala des Situational Judgement Tests (SEC-SJT), bei fehlenden Antworten nicht ausgewertet werden.

2.3.4. Auswertung von Hand

2.3.4.1. Berechnung der Skalenrohwerte

Für die Auswertung wird das Ergebnis auf den Globalskalen wie auch den einzelnen Subskalen zunächst in Form von Rohwerten berechnet und anschließend standardisiert. Vor der Berechnung der Rohwerte sollte die Qualität der Bearbeitung geprüft werden (siehe Abschnitt *Prüfung der Qualität der Bearbeitung*, S. 35). Bestehen Zweifel an der Qualität

Auswertung

der Bearbeitung, sollte die Auswertung und damit auch eine spätere Interpretation des Ergebnisses nur unter Vorbehalt erfolgen.

Um den Rohwert für eine einzelne Skala zu ermitteln, werden die Antworten auf die Items der entsprechenden Skala zu einem Summenwert addiert. Da alle Items in Richtung der Merkmalsausprägungen formuliert sind, können die numerischen Werte der Antworten einfach addiert werden. Alle Items sind auf einer fünfstufigen Likert-Skala zu beantworten, auf der jeder Wert der Likert-Skala verbal verankert ist. Die verbalen Anker entsprechen den folgenden numerischen Werten:

- ◆ *nie* = 1,
- ◆ *selten* = 2,
- ◆ *gelegentlich* = 3,
- ◆ *häufig* = 4,
- ◆ *immer* = 5.

Auf dem Fragebogen zur Selbsteinschätzung und dem Fragebogen zur Fremdeinschätzung finden sich keine Hinweise darauf, welche Items welcher Skala zuzuordnen sind. Die Items sind fortlaufend nummeriert und so randomisiert, dass Items derselben Skala maximalen Abstand zueinander aufweisen. Die Zuordnung ergibt sich aus der Auswertungstabelle auf dem entsprechenden Auswertungsbogen bzw. kann auch dem Itemschlüssel in Tabelle A8 des Anhangs (siehe S. 149) entnommen werden.

Für die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) ist es möglich, bei bis zu maximal vier fehlenden Antworten (insgesamt unter 10.0% fehlende Antworten) mittels Interpolation einen geschätzten Rohwert zu berechnen (siehe Abschnitt *Interpolation des Rohwerts der Globalskala des Fragebogens zur Selbsteinschätzung oder der Globalskala des Fragebogens zur Fremdeinschätzung*, S. 42).

Bei fehlenden Antworten im Situational Judgement Test, auf den Subskalen des Fragebogens zur Selbsteinschätzung oder auf den Subskalen des Fragebogens zur Fremdeinschätzung sollte von einer Auswertung der von fehlenden Antworten betroffenen Skalen abgesehen werden. Da die Subskalen des Fragebogens zur Selbsteinschätzung, die Subskalen des Fragebogens zur Fremdeinschätzung und die Globalskala des Situational Judgement Tests (SEC-SJT) nur relativ wenige Items enthalten, wäre eine Schätzung des Rohwerts für die jeweiligen Skalen erheblich fehleranfälliger als die Schätzung des Rohwerts für die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) oder die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) mit jeweils 41 Items. Alternativ kann noch versucht werden, fehlende Antworten im Nachhinein von der Testperson bzw. der einschätzenden Person ergänzen zu lassen (siehe Abschnitt *Fehlende Antworten*, S. 40).

Um die Auswertung zu erleichtern, wurde darauf verzichtet, die Subskalen entsprechend ihrer Itemanzahl zu gewichten. Zwar fließen dadurch die Subskalen zu unterschiedlichen Anteilen in die Globalskala ein, aufgrund der Homogenität der Subskalen ist dieser Effekt jedoch vernachlässigbar.

2.3.4.2. Interpolation des Rohwerts der Globalskala des Fragebogens zur Selbsteinschätzung oder der Globalskala des Fragebogens zur Fremdeinschätzung

Wie bereits angesprochen, kann der Rohwert der Globalskalen des Fragebogens zur Selbsteinschätzung (SEC-SR) oder der Rohwert der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) bei einer bis vier fehlenden Antwort(en) mittels Interpolation geschätzt werden. Eine Schätzung mittels Interpolation bei fehlenden Antworten ist nur für die Globalskalen SEC-SR oder SEC-OR möglich. Die Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung und die Globalskala des Situational Judgement Tests können bei fehlenden Antworten nicht ausgewertet werden.

Formel zur Schätzung des Rohwerts der Globalskalen SEC-SR oder SEC-OR bei maximal vier unbeantworteten Items:

$$\frac{\text{Summenwert aller beantworteten Items}}{\text{Anzahl aller beantworteten Items}} \cdot 41$$

Das Ergebnis der Interpolation wird dann ganzzahlig ab- (Nachkommastellen < 0.50) bzw. aufgerundet (Nachkommastellen ≥ 0.50). Die Formel sollte jedoch ausschließlich bei nicht mehr als vier unbeantworteten Items zur Schätzung des Rohwerts der Globalskala SEC-SR oder SEC-OR verwendet werden, da die Schätzung mit steigender Anzahl unbeantworteter Items immer fehleranfälliger wird. Wenn der Wert für die Globalskala SEC-SR oder SEC-OR mittels Interpolation für eine Testperson geschätzt wurde, sollte dies bei der Interpretation des Ergebnisses beachtet werden. Interpolierte Rohwerte sollten daher stets durch die Markierung mit einem * kenntlich gemacht werden.

Die hier vorgeschlagene Art der Interpolation, wie auch andere Strategien bei fehlenden Antworten, sind aus testtheoretischer Perspektive nicht optimal, denn beispielsweise könnte fehlenden Antworten auch eine systematische Ursache zu Grunde liegen (eingeschränkte Selbstreflexion oder mangelnde Antwortbereitschaft für die unbeantworteten

Items oder das nicht eingeschätzte Merkmal). Nichtsdestotrotz besteht durch die Interpolation eine gute Möglichkeit, zumindest noch die Globalskalen SEC-SR und SEC-OR auszuwerten zu können, falls es nicht mehr möglich ist, fehlende Antworten noch im Nachhinein einzuholen.

2.3.5. Auswertung mittels Auswertungsbögen

Zur Beschreibung der Auswertung wurde eine Person mit einem unauffälligen Testergebnis aus der Gesamtstichprobe (siehe Abschnitt *Beschreibung der Gesamtstichprobe*, S. 120) ausgewählt. Zur Wahrung ihrer Anonymität werden keine Angaben zu Alter und Geschlecht gemacht. Ein Teil der Auswertung betrifft bereits grundlegende Entscheidungen der Interpretation (Wahl der geeigneten Norm, Wahl der Vertrauenswahrscheinlichkeit). Daher wird die jeweilige Beschreibung der Auswertung im Abschnitt *Interpretation mit dem zugehörigen Profilbogen* (siehe S. 76) fortgesetzt, nachdem im Abschnitt zur Interpretation wichtige *Grundlagen der Interpretation psychologischer Testergebnisse* (siehe S. 53) angeschnitten wurden. Grundlegenden Entscheidungen der Interpretation sollten nur von Fachpersonal mit psychologischer Ausbildung und Erfahrung in psychologischer Diagnostik getroffen werden.

2.3.5.1. Auswertung des Fragebogens zur Selbsteinschätzung

Für die Auswertung wird der *SEC-I-SR Auswertungsbogen* und der Profilbogen benötigt. Im Profilbogen können die Testergebnisse aus dem Fragebogen zur Selbsteinschätzung, dem Fragebogen zur Fremdeinschätzung und dem Situational Judgement Test gemeinsam dokumentiert werden, es können jedoch beispielsweise auch nur Testergebnisse aus der Erhebung der Selbsteinschätzung dokumentiert werden. Die für die Auswertung erforderlichen Schritte sind auf dem Auswertungsbogen zusätzlich in kurzen Stichpunkten aufgeführt.

Vor der Auswertung: Prüfung der Qualität der Bearbeitung

Bevor mit der Auswertung begonnen wird, wird die Qualität der Bearbeitung geprüft. Wurden von der Testleitung auf der letzten Seite des Fragebogens zur Selbsteinschätzung keine Beobachtungen dokumentiert und zeigt die Selbsteinschätzung keine Auffälligkeiten im Antwortverhalten der Testperson (siehe Abschnitt *Prüfung der Qualität der Bearbeitung*, S. 35), kann mit der Auswertung begonnen werden.

Schritt 1: Eintragen der numerischen Werte

Für die Auswertung von Hand werden die numerischen Werte der Antworten auf die einzelnen Items in die Auswertungstabelle des *SEC-I-SR Auswertungsbogens* eingetragen. Schritt 1 ist in Abbildung 6 dargestellt.

Abbildung 6

Schritt 1 der Auswertung des Fragebogens zur Selbsteinschätzung (SEC-I-SR) mittels SEC-I-SR Auswertungsbogen

SEC-I-SR

Bitte kreuze für jede Aussage die Abstufung an, die dich am besten beschreibt.

	nie	selten	gelegentlich	häufig	immer
1) Ich versuche zu helfen, wenn andere ängstlich sind.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
2) Ich kann andere beruhigen, wenn sie ängstlich sind.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
3) Ich bemerke, wenn ich stolz bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
4) Ich kann mich selbst beruhigen, wenn ich angespannt bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
5) Ich will mich gleich nach einem Streit wieder vertragen.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
6) Ich bemerke, wenn andere schlecht gelaunt sind.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
7) Ich kann verstehen, weshalb andere angespannt sind.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8) Ich bedanke mich, wenn andere etwas für mich tun.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
9) Ich unterstütze andere.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
10) Ich versuche anderen zu helfen, wenn ich das Gefühl habe, dass sie Hilfe brauchen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
11) Ich kann andere beruhigen, wenn sie angespannt sind.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12) Ich bemerke, wenn ich traurig bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
13) Ich kann mich selbst beruhigen, wenn ich ängstlich bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
14) Es beschäftigt mich, wenn ich Streit mit anderen habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

SEC-I-SR AUSWERTUNGSBOGEN

Auswertungsschritte

- Tragen Sie die numerischen Werte der Antworten auf die einzelnen Items der Selbsteinschätzung in die dafür vorgesehenen Zellen (die jeweils hell hinterlegte Zelle) der Auswertungstabelle ein. Die verbalen Anker entsprechen dabei den folgenden numerischen Werten: nie = 1, selten = 2, gelegentlich = 3, häufig = 4, immer = 5.
- Bilden Sie die Spaltensummen auf dieser Seite und übertragen Sie diese in die oberste Zeile der Tabellenfortsetzung auf der nächsten Seite.
- Bilden Sie für die Tabellenfortsetzung die Spaltensummen. Diese entsprechen den Rohwerten der Subskalen EBE-SR, ECO-SR, PEA-SR, PES-SR, REA-SR, RES-SR, SOB-SR und SON-SR.
- Bilden Sie die Zeilensumme (Summe der Spaltensummen). Diese entspricht dem Rohwert der Globalskala SEC-SR.

Auswertungstabelle

Seite 4	Skala	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	SEC-SR
Item										
1		3								
2						3				
3							5			
4									4	
5									3	
6										
7			2		4					
8										5
9									3	
10		4								
11						2				
12										
13							4			
14								4		
Zwischen-summe										

Fortsetzung auf der folgenden Seite.

Erasmus+ 4 von 7

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

In der Auswertungstabelle ist für jedes Item eine Zeile und für jede Skala eine Spalte vorhanden. Da jedes Item nur einer Skala zugeordnet ist, ist in jeder Zeile auch nur für eine Spalte eine hell hinterlegte Zelle vorhanden, in die der numerische Wert eingetragen werden kann. In die dunkel hinterlegten Zellen sind keine Werte einzutragen. Beispielsweise ist Item 1 des Fragebogens zur Selbsteinschätzung in der Auswertungstabelle des *SEC-I-SR Auswertungsbogens* der Subskala EBE-SR zugeordnet, daher befindet sich in der Zeile für Item 1 die hell hinterlegte Zelle, in die der numerische Wert des Items eingetragen wird, in der Spalte für die Subskala EBE-SR. In der Zeile für Item 2 befindet sich die hell hinterlegte Zelle in der Spalte für die Subskala REA-SR, Item 2 des Fragebogens zur

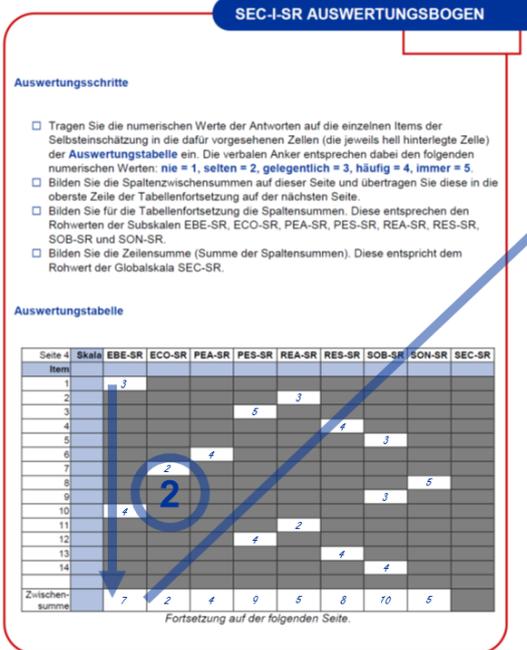
Auswertung

Selbsteinschätzung ist also der Subskala REA-SR zugeordnet und so weiter. Die numerischen Werte ergeben sich aus den von der Testperson auf der Likert-Skala angekreuzten verbalen Anker, die den folgenden numerischen Werten entsprechen:

- ◆ *nie* = 1,
- ◆ *selten* = 2,
- ◆ *gelegentlich* = 3,
- ◆ *häufig* = 4,
- ◆ *immer* = 5.

Abbildung 7

Schritte 2, 3 und 4 der Auswertung des Fragebogens zur Selbsteinschätzung (SEC-I-SR) mittels SEC-I-SR Auswertungsbogen



SEC-I-SR AUSWERTUNGSBOGEN

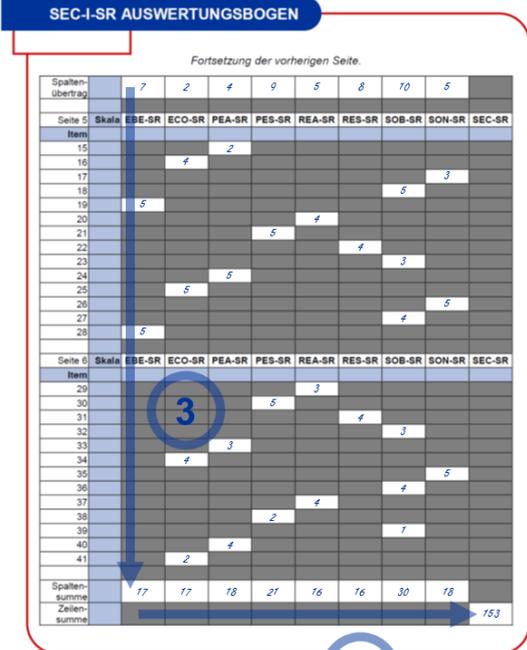
Auswertungsschritte

- Tragen Sie die numerischen Werte der Antworten auf die einzelnen Items der Selbsteinschätzung in die dafür vorgesehenen Zellen (die jeweils hell hinterlegte Zelle) der **Auswertungstabelle** ein. Die verbalen Anker entsprechen dabei den folgenden numerischen Werten: **nie** = 1, **selten** = 2, **gelegentlich** = 3, **häufig** = 4, **immer** = 5.
- Bilden Sie die Spaltenzwischensummen auf dieser Seite und übertragen Sie diese in die oberste Zeile der Tabellenfortsetzung auf der nächsten Seite.
- Bilden Sie für die Tabellenfortsetzung die Spaltensummen. Diese entsprechen den Rohwerten der Subskalen EBE-SR, ECO-SR, PEA-SR, PES-SR, REA-SR, RES-SR, SOB-SR und SON-SR.
- Bilden Sie die Zeilensumme (Summe der Spaltensummen). Diese entspricht dem Rohwert der Globalskala SEC-SR.

Auswertungstabelle

Seite 4	Skala	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	SEC-SR
Item										
1		3								
2						3				
3					5					
4							4			
5								3		
6					4					
7									5	
8										5
9										5
10									3	
11										5
12						4				
13									4	
14										4
Zwischen-summe		7	2	4	9	5	8	10	5	

Fortsetzung auf der folgenden Seite.



SEC-I-SR AUSWERTUNGSBOGEN

Fortsetzung der vorherigen Seite.

Spalten-übertrag	7	2	4	9	5	8	10	5		
Seite 5	Skala	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	SEC-SR
Item										
15				2						
16										
17				4						
18										
19									5	
20						4				
21					5					
22								4		
23									3	
24										
25				5	5					
26										5
27									4	
28										5
Seite 6	Skala	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	SEC-SR
Item										
29										
30										
31						5		3		
32								4		
33									3	
34										
35						4				5
36									4	
37										5
38										
39										7
40										
41						2		4		
Zwischen-summe		17	17	18	21	16	16	30	18	
Zeilensumme										153

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

Schritt 2: Zwischenschritt zur Berechnung der Rohwerte der Subskalen

Wurden alle numerischen Werte in die entsprechenden hell hinterlegten Zellen eingetragen, werden auf Seite 3 die Spaltenzwischensummen gebildet und in die Zeile *Zwischen-summe* eingetragen. Die Spaltenzwischensummen werden dann auf Seite 4 in die oberste

Zeile der Tabellenfortsetzung (*Spaltenübertrag*) übertragen. Schritt 2 ist in Abbildung 7 dargestellt.

Schritt 3: Berechnung der Rohwerte der Subskalen

Als nächstes werden auf Seite 4 die Spaltensummen gebildet und in die Zeile *Spaltensumme* eingetragen. Die Spaltensummen werden jeweils über die gesamte Spalte gebildet, der Spaltenübertrag fließt wie jede andere Zeile auch in die Spaltensummen ein. Schritt 3 ist in Abbildung 7 dargestellt. Die Spaltensummen entsprechen den Rohwerten der Testperson auf den Subskalen.

Abbildung 8

Schritt 5 der Auswertung des Fragebogens zur Selbsteinschätzung (SEC-I-SR) mittels SEC-I-SR Auswertungsbogen und Profilbogen

SEC-I-SR AUSWERTUNGSBOGEN

Fortsetzung der vorherigen Seite.

Spaltenübertrag	7	2	4	9	5	8	10	5	
Seite 0 Skala	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	SEC-SR
Item 15			2						
Item 16		4							
Item 17								3	
Item 18							5		
Item 19	5								
Item 20					4				
Item 21				5					
Item 22						4			
Item 23							3		
Item 24			5	5					
Item 25								5	
Item 26									5
Item 27								4	
Item 28	5								
Seite 0 Skala	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	SEC-SR
Item 29					3				
Item 30					5				
Item 31						4			
Item 32							3		
Item 33			3						
Item 34		4							
Item 35								5	
Item 36							4		
Item 37					4				
Item 38				2					
Item 39								1	
Item 40									
Item 41		2	4						
Spaltensumme	17	17	18	21	16	16	30	18	
Zeilensumme									
	153								

PROFILBOGEN

Normierungsschritte

- Übertragen Sie die Rohwerte (Summen aus dem SEC-I-SR, SEC-I-OR und SEC-SJT Auswertungsbogen) der jeweiligen Skalen (Subskalen und Globalskalen) in die dafür vorgesehenen Zellen der **Ergebnistabelle**.
- Wählen Sie eine geeignete Norm (Gesamtstichprobe, Stichprobe weiblichen bzw. männlichen Geschlechts) und entnehmen Sie dann die zugehörigen T-Werte bzw. Prozentränge (PR) dem Testmanual.

Ergebnistabelle

Norm: Gesamt weiblich männlich

	Rohwert	T-Wert	PR		Rohwert	T-Wert	PR
Skala				Skala			
SEC-SJT				SEC-OR			
SEC-SR	153			EBE-OR			
EBE-SR	17			ECO-OR			
ECO-SR	17			PEA-OR			
PEA-SR	18			PES-OR			
PES-SR	21			REA-OR			
REA-SR	16			RES-OR			
RES-SR	16			SOB-OR			
SOB-SR	30			SON-OR			
SON-SR	18						

5



4 von 4



3 von 5

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

Schritt 4: Berechnung des Rohwerts der Globalskala

Abschließend wird auf Seite 4 die Zeilensumme über die Spaltensummen gebildet, es werden also die Rohwerte der Subskalen addiert, und in die hell hinterlegte Zelle der Zeile *Zeilensumme* eingetragen. Schritt 4 ist in Abbildung 7 dargestellt. Die Zeilensumme entspricht dem Rohwert der Testperson auf der Globalskala SEC-SR.

Schritt 5: Übertragen der Rohwerte der Skalen

Die berechneten Rohwerte der Globalskala und der Subskalen werden aus dem *SEC-I-SR Auswertungsbogen* in die Ergebnistabelle auf Seite 3 des Profilbogens übertragen. Schritt 5 ist in Abbildung 8 dargestellt.

2.3.5.2. Auswertung des Fragebogens zur Fremdeinschätzung

Für die Auswertung wird der *SEC-I-OR Auswertungsbogen* und der Profilbogen benötigt. Da der Fragebogen zur Fremdeinschätzung parallel zum Fragebogen zur Selbsteinschätzung konstruiert wurde (siehe Abschnitt *Fragebogen zur Fremdeinschätzung*, S. 14), erfolgt die Auswertung des Fragebogens zur Fremdeinschätzung analog zur Auswertung des Fragebogens zur Selbsteinschätzung (siehe Abschnitt *Auswertung des Fragebogens zur Selbsteinschätzung*, S. 43).

2.3.5.3. Auswertung des Situational Judgement Test

Anders als der Fragebogen zur Selbsteinschätzung oder der Fragebogen zur Fremdeinschätzung, die Merkmale auch auf Subskalen erfassen, erfasst der Situational Judgement Test nur die Globalskala SEC-SJT. Für die Auswertung wird der *SEC-SJT Auswertungsbogen* und der Profilbogen benötigt. Genau wie auf dem *SEC-I-SR Auswertungsbogen* und dem *SEC-I-OR Auswertungsbogen* werden auf dem *SEC-SJT Auswertungsbogen* die für die Auswertung erforderlichen Schritte noch einmal in kurzen Stichpunkten aufgeführt.

Vor der Auswertung: Prüfung der Qualität der Bearbeitung

Auch für den Situational Judgement Test gilt, dass vor der Auswertung die Qualität der Bearbeitung geprüft werden sollte (siehe Abschnitt *Prüfung der Qualität der Bearbeitung*, S. 35). Wurden von der Testleitung auf der letzten Seite des Situational Judgement Tests keine Beobachtungen dokumentiert und zeigen sich keine anderen Auffälligkeiten kann mit der Auswertung begonnen werden.

Schritt 1: Eintragen der numerischen Werte

Im Gegensatz zur Auswertungstabelle auf dem *SEC-I-SR Auswertungsbogen* und dem *SEC-I-OR Auswertungsbogen* markieren die hell hinterlegten Zellen der Auswertungstabelle auf dem *SEC-SJT Auswertungsbogen* nicht die Zuordnung eines Items zu einer Skala, sondern die jeweilige Antwort nach Konsensus- bzw. Expertenmethode.

Die Antworten auf die Items des Situational Judgement Tests werden in die Auswertungstabelle übertragen, indem eine 1 in die dafür vorgesehene hell hinterlegte Zelle der Auswertungstabelle des *SEC-SJT Auswertungsbogens* eingetragen wird, wenn die entsprechende Antwortalternative gewählt wurde. Wurde die entsprechende Antwortalternative nicht gewählt, wird eine 0 in die dafür vorgesehene hell hinterlegte Zelle eingetragen. Schritt 1 ist in Abbildung 9 dargestellt.

Abbildung 9

Schritte 1 und 2 der Auswertung des Situational Judgement Tests (SEC-SJT) mittels SEC-SJT Auswertungsbogen

SEC-SJT

Was ist für dich die beste Antwort? Bitte kreuze die für dich beste Antwort an. Findest du keine Antwort gut? Dann kreuze einfach die Antwort an, die für dich noch am besten ist.

5) Der Trainer kritisiert Max häufig. Max findet das unfair. Was sollte Max tun, damit es ihm besser geht?

- Abwarten. Der Trainer wird schon irgendwann aufhören, Max zu kritisieren.
- Sich verteidigen. Niemand sollte so oft kritisiert werden.
- Sich eine andere Mannschaft suchen, um den Trainer nicht mehr zu sehen.
- Seinem Trainer sagen: „Ich fühle mich unfair behandelt.“
- Seinem Trainer sagen: „Wenn ich weiter kritisiert werde, suche ich mir eine andere Mannschaft.“

6) Ulla weiß, dass ihr Arbeitskollege Michael sehr oft traurig ist. Heute scheint Michael besonders traurig zu sein. Was sollte Ulla tun, damit es Michael besser geht?

- Nichts. Michael wird es irgendwann von alleine wieder besser gehen.
- Mit ihrem Chef darüber sprechen. Vielleicht kann er Michael helfen.
- Michael fragen: „Ist alles in Ordnung? Kann ich irgendetwas für dich tun?“
- Michael zum Lachen bringen.
- Michael fragen: „Wieso bist du denn jetzt schon wieder traurig?“

SEC-SJT AUSWERTUNGSBOGEN

Auswertungsschritte

- Übertragen Sie die Antworten auf die einzelnen Items des Situational Judgement Test, indem Sie eine 1, wenn die entsprechende Antwortalternative gewählt wurde, bzw. eine 0, wenn die entsprechende Antwortalternative nicht gewählt wurde, in die dafür vorgesehenen Zellen (die jeweils hell hinterlegte Zelle) der Auswertungstabelle eintragen.
- Bilden Sie die Summe aller derart übertragenen Antworten. Diese entspricht dem Rohwert der Globalskala SEC-SJT (mit einem maximalen Rohwert von 10).

Auswertungstabelle

	Antwort-alternative	1	2	3	4	5	Skala	SEC-SJT
Item								
1				1				
2					1			
3			1					
4		1						
5					0			
6				1				
7					1			
8					0			
9			1					
10					1			
Summe								8

Erasmus+ 6 von 9 3 von 3 Erasmus+

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

Für jedes Item gibt es eine Antwort nach Konsensus- bzw. Expertenmethode, daher ist in jeder Zeile eines Items auch nur für eine Antwortalternative, also nur in einer der Spalten,

Auswertung

eine hell hinterlegte Zelle vorhanden. Beispielsweise ist für Item 1 die dritte Antwortalternative die Antwort nach Konsensus- bzw. Expertenmethode (zu erkennen an der hell hinterlegten Zelle in der Spalte Antwortalternative 3), für Item 2 ist die vierte Antwortalternative die Antwort nach Konsensus- bzw. Expertenmethode (zu erkennen an der hell hinterlegten Zelle in der Spalte Antwortalternative 4) und so weiter.

Schritt 2: Berechnung des Rohwerts auf der Globalskala

Wurden alle Antworten auf die zuvor beschriebene Weise übertragen, werden die numerischen Werte aller Antworten addiert und die Summe in die Zeile *Summe* eingetragen. Der dabei maximal erreichbare Rohwert beträgt 10. Schritt 2 ist in Abbildung 9 dargestellt. Die Summe entspricht dem Rohwert der Testperson auf der Globalskala SEC-SJT.

Abbildung 10

Schritt 3 der Auswertung des Situational Judgement Tests SEC-SJT mittels SEC-SJT Auswertungsbogen und Profilbogen

SEC4VET
Erasmus-Projekt
Social Emotional Competence for VET

SEC-SJT AUSWERTUNGSBOGEN

Auswertungsschritte

- Übertragen Sie die Antworten auf die einzelnen Items des Situational Judgement Test, indem Sie eine **1**, wenn die entsprechende Antwortalternative gewählt wurde, bzw. eine **0**, wenn die entsprechende Antwortalternative nicht gewählt wurde, in die dafür vorgesehenen Zellen (die jeweils hell hinterlegte Zelle) der **Auswertungstabelle** eintragen.
- Bilden Sie die Summe aller derart übertragenen Antworten. Diese entspricht dem Rohwert der Globalskala SEC-SJT (mit einem maximalen Rohwert von 10).

Auswertungstabelle

	Antwort- alternative	1	2	3	4	5	Skala	SEC-SJT
Item								
1				1				
2					1			
3		1						
4		1						
5						0		
6				1				
7				1				
8					0			
9			1					
10				1				
Summe								8

SEC4VET
Erasmus-Projekt
Social Emotional Competence for VET

PROFILBOGEN

Normierungsschritte

- Übertragen Sie die Rohwerte (Summen aus dem SEC-I-SR, SEC-I-OR und SEC-SJT Auswertungsbogen) der jeweiligen Skalen (Subskalen und Globalskalen) in die dafür vorgesehenen Zellen der **Ergebnistabelle**.
- Wählen Sie eine geeignete Norm (Gesamtstichprobe, Stichprobe weiblichen bzw. männlichen Geschlechts) und entnehmen Sie dann die zugehörigen T-Werte bzw. Prozentränge (PR) dem Testmanual.

Ergebnistabelle

Norm: Gesamt weiblich männlich

	Rohwert	T-Wert	PR		Rohwert	T-Wert	PR
Skala				Skala			
SEC-SJT	8			SEC-OR			
SEC-SR				EBE-OR			
EBE-SR				ECO-OR			
ECO-SR				PEA-OR			
PEA-SR				PES-OR			
PES-SR				REA-OR			
REA-SR				RES-OR			
RES-SR				SOB-OR			
SOB-SR				SON-OR			
SON-SR							

3

3 von 3

3 von 5

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

Schritt 3: Übertragen des Rohwerts der Skala

Abschließend wird der Rohwert der Globalskala SEC-SJT aus dem *SEC-SJT Auswertungsbogen* in die Ergebnistabelle auf Seite 3 des Profilbogens übertragen. Schritt 3 ist in Abbildung 10 dargestellt.

2.3.6. Computerbasierte Auswertung

Für die computerbasierte Auswertung wird eine Auswertungsdatei (Microsoft® Excel® Arbeitsmappe) zur Verfügung gestellt. Für die Auswertung werden die Antworten aus dem Fragebogen zur Selbsteinschätzung, dem Fragebogen zur Fremdeinschätzung bzw. dem Situational Judgement Test der Reihenfolge nach in die Auswertungsdatei übertragen, die Auswertung erfolgt dann vollständig automatisiert. Die Auswertungsdatei enthält den *SEC-I-SR Auswertungsbogen*, den *SEC-I-OR Auswertungsbogen*, den *SEC-SJT Auswertungsbogen* und den Profilbogen, die mit dem entsprechenden Papierbogen übereinstimmen und separat gedruckt werden können. Wurde die Auswertung mit Hilfe dieses Testmanuals und der Testmaterialien erarbeitet, ist die Auswertungsdatei selbsterklärend. Da die Auswertung von Hand sehr ausführlich erläutert wird (siehe Abschnitt *Auswertung von Hand*, S. 40; und Abschnitt *Auswertung mittels Auswertungsbögen*, S. 43), wird nicht weiter auf die computerbasierte Auswertung eingegangen. Die computerbasierte Auswertung ist deutlich ökonomischer und weniger anfällig für Fehler und sollte daher, wenn möglich, der Auswertung von Hand vorgezogen werden.

Auch in der computerbasierten Auswertung wird der Rohwert der Globalskalen SEC-SR oder SEC-OR bei bis zu maximal vier fehlenden Antworten interpoliert. Wurde der Wert automatisch interpoliert, ist der interpolierte Rohwert an der Markierung mit einem * zu erkennen.

2.4. Interpretation

2.4.1. Qualifikation der interpretierenden Person

Anders als Erhebung und Auswertung, die auch von gut geschulten psychologischen Hilfskräften durchgeführt werden können (siehe Abschnitt *Durchführung*, S. 29; und Abschnitt *Auswertung*, S. 34), sollten die Testergebnisse nur von Fachpersonal mit psychologischer Ausbildung interpretiert werden. Die interpretierende Person sollte über fundierte Fachkenntnisse und Erfahrung in psychologischer Diagnostik verfügen, in der Fachliteratur zu sozialen und emotionalen Kompetenzen belesen und mit sämtlichen Inhalten dieses Testmanuals vertraut sein.

Der Umgang mit den diagnostischen Verfahren kann im *Workshop zur Diagnostik sozial-emotionaler Kompetenzen* (siehe S. 25) geschult werden, eine psychologische Ausbildung kann der Workshop jedoch nicht ersetzen. Wird die Schulung von Fachpersonal mit der entsprechenden Qualifikation durchgeführt, können zwar wichtige Grundlagen der Interpretation von Ergebnissen psychologischer Tests interaktiv vermittelt bzw. aufgefrischt und Unsicherheiten abgebaut werden, nichtsdestotrotz sollte die Interpretation qualifiziertem Fachpersonal vorbehalten bleiben.

2.4.2. Allgemeine Hinweise zur Interpretation

Bevor Testergebnisse überhaupt interpretiert werden, sollte sichergestellt worden sein, dass die Qualität der Bearbeitung des Fragebogens im Rahmen der Auswertung geprüft wurde (siehe Abschnitt *Prüfung der Qualität der Bearbeitung*, S. 35) und Subskalen des Fragebogens zur Selbsteinschätzung des Fragebogens zur Fremdeinschätzung sowie die Globalskala des Situational Judgement Tests bei fehlenden Antworten nicht ausgewertet wurden. Von der Interpretation sollte abgesehen werden, wenn Zweifel an der Qualität der Bearbeitung bestehen.

Manchmal wird dazu geraten, bei der Interpretation von Testergebnissen auf einzelne Items zurückzugreifen und ihre Beantwortung ebenfalls zu interpretieren. Von dieser Praxis raten wir jedoch ausdrücklich ab. Die von Skalen erfassten Merkmale werden auch als latente Konstrukte bezeichnet. Latente Konstrukte sind der unmittelbaren Beobachtung nicht zugänglich und werden erst durch ihre Operationalisierung messbar. Die Operationalisierung ist die Messbarmachung eines latenten Konstrukts durch die Beschreibung unmittelbar beobachtbarer Tatsachengegenstände. Im Fall psychologischer Tests sind diese Tatsachengegenstände Items. Items beschreiben beobachtbares (manifestes) Verhalten, von dem angenommen wird, dass dieses von dem dahinterliegenden – nicht unmittelbar beobachtbaren (latenten) – Konstrukt beeinflusst wird bzw. dass das latente Konstrukt dessen Ursache ist. Die in einer Skala enthalten Items sind immer nur eine Teilmenge aller möglichen Items, die für die Operationalisierung eines latenten Konstrukts herangezogen werden können. Demzufolge misst ein Fragebogen ausschließlich latente Konstrukte und nicht Reaktionen auf einzelne Items.

Im Allgemeinen messen Fragebögen zur Selbst- und Fremdeinschätzung auch nicht objektive Tatsachen, sondern wie sich eine Testperson selbst einschätzt bzw. durch eine oder gegebenenfalls mehrere andere Personen eingeschätzt wird. Selbst- und Fremdeinschätzungen sind damit subjektiver Art und beide sind von Fähigkeiten (beispielsweise der Fähigkeit zur Selbstreflexion oder der Fähigkeit zur nicht wertenden Beobachtung), Motiven (beispielsweise Täuschung oder Zeiteinsparung), Einstellungen (beispielsweise Moti-

vation oder Sympathie), Tageseinflüssen und weiteren Faktoren abhängig. Werden Fremdeinschätzungen derselben Testperson durch verschiedene andere Testpersonen erhoben und integriert, können Fremdeinschätzungen objektiviert werden. Insbesondere bei innerem emotionalen Erleben bleiben jedoch verschiedene Aspekte dieses Erlebens Fremdeinschätzungen gegenüber verborgen, die nur durch Selbsteinschätzungen erschlossen werden können.

Es sollte daher selbstverständlich sein, dass jede Form von Einschätzung – unabhängig davon, ob es sich um Selbst- oder Fremdeinschätzungen handelt – meist nur eine eingeschränkte Objektivität besitzt. Daher sollten diagnostische Entscheidungen generell, aber insbesondere solche mit weitreichenden Konsequenzen, niemals allein auf Basis des Testergebnisses aus einer Selbst- oder einer Fremdeinschätzung getroffen werden. Um die Aussagekraft von Einschätzungen zu erhöhen, sollten stets zusätzliche Informationen erschlossen und integriert werden. Wie zuvor erwähnt, können Selbsteinschätzungen durch Fremdeinschätzung und umgekehrt Fremdeinschätzungen durch Selbsteinschätzungen ergänzt werden. Es können weitere Fremdeinschätzungen derselben Testperson durch weitere andere Personen erhoben werden. Zusätzlich können auch, sofern die Möglichkeit dazu besteht, objektive Leistungsdaten erhoben oder ergänzend biografische Daten eingeholt werden. Mit dem Fragebogen zur Selbsteinschätzung, dem Fragebogen zur Fremdeinschätzung sowie dem Situational Judgement Test steht für die Erschließung diagnostischer Informationen eine breite Palette an Möglichkeiten zur Verfügung.

In erster Linie sollte die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR), die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) und die Globalskala des Situational Judgement Tests (SEC-SJT) interpretiert werden. Aufgrund der geringer ausfallenden Reliabilitäten (siehe Abschnitt *Konsistenzanalyse*, S. 94) und der Homogenität der Subfacetten bzw. der relativ hohen Interkorrelationen der Subskalen (siehe Abschnitt *Skaleninterkorrelationen*, S. 93; und Abschnitt *Multitrait-Multimethod-Analyse*, S. 101) sollten die Subskalen vorsichtiger anhand der in den Profilbogen (siehe Abschnitt *Interpretation einer Profillinie*, S. 61; und Abschnitt *Profilreliabilität*, S. 96) eingezeichneten standardisierten Werte und Vertrauensbereiche interpretiert werden.

Für die Interpretation sollten ausschließlich standardisierte Werte herangezogen werden (siehe Abschnitt *Standardisierte Werte*, S. 53). Standardisierte Werte psychologischer Tests sind in der Regel keine Trennwerte, ab denen einer Testperson ein bestimmtes Merkmal fest zugeschrieben werden könnte, denn sie sind nicht frei von Messfehlern und können daher mehr oder weniger vom wahren Wert einer Testperson abweichen. Standardisierte Werte sollten daher nur unter Einbezug der Reliabilität interpretiert werden (siehe Abschnitt *Messfehler und Reliabilität*, S. 55; und Abschnitt *Reliabilität*, S. 94).

Bei der Interpretation intraindividuelle Differenzen (siehe Abschnitt *Interpretation von Differenzen*, S. 63) ist zu berücksichtigen, dass die Stabilitäten der Skalen nicht empirisch

überprüft wurden. Es ist möglich, dass intraindividuelle Differenzen innerhalb der Skalen zeitlich nicht stabil sind und nur zufällige (unsystematische) oder tagesformbedingte (systematische) Schwankungen abbilden. Um dem zu begegnen sollte zumindest für die Wahl einer geeigneten Vertrauenswahrscheinlichkeit möglichst konservativ vorgegangen werden.

2.4.3. Grundlagen der Interpretation psychologischer Testergebnisse

Im folgenden Abschnitt werden wichtige fachliche Grundlagen der Interpretation psychologischer Testergebnisse angesprochen und gängige Fachtermini näher erläutert. Fachpersonal mit fundierten Kenntnissen im Bereich der psychologischen Diagnostik wird hier wenig Neues finden. Bestehen jedoch Unsicherheiten im Bereich der psychologischen Diagnostik liefert der Abschnitt eine Auffrischung der wichtigsten Grundlagen.

2.4.3.1. Standardisierte Werte

In psychologischen Tests wird der Rohwert einer Testperson meist durch einfaches Auszählen der Antworten ermittelt. Doch bei Rohwerten ist die Metrik unbekannt und damit ist auch nicht klar, welche Bedeutung das Testergebnis einer Testperson hat. Um die Bedeutung des Rohwerts einschätzen zu können, ist es erforderlich, diesen in Bezug zu den Testergebnissen weiterer Personen zu setzen. Die Testergebnisse weiterer Personen werden hierfür in Vergleichsgruppen, so genannten Normen, zusammengefasst. Das Testergebnis einer Testperson kann dann in Relation zu den Testergebnissen der Norm gesetzt werden. Dadurch wird der Rohwert in einen standardisierten Wert überführt, erhält eine Metrik und kann dann auch interpretiert werden.

Bei standardisierten Werten ist unmittelbar ersichtlich, ob ein entsprechender Wert im unterdurchschnittlichen, durchschnittlichen oder überdurchschnittlichen Bereich liegt. Die Interpretation der Ergebnisse psychologischer Tests sollte ausschließlich anhand der standardisierten Werte und niemals anhand der Rohwerte erfolgen. Die für die Standardisierung herangezogene Vergleichsgruppe (Norm) sollte dabei stets berücksichtigt werden. Beispielsweise ist es durchaus möglich, dass dasselbe Testergebnis in Rohwerten in Abhängigkeit von der herangezogenen Norm entweder im überdurchschnittlichen oder im durchschnittlichen Bereich liegen kann.

Analogie zur Veranschaulichung

Stellen Sie sich vor, Sie würden das Körpergewicht einer Person kennen, wüssten allerdings überhaupt nichts über das Körpergewicht anderer Personen. Sie hätten also keinerlei Bezugsrahmen für die Einordnung des Gewichts dieser Person. Sie könnten nicht sagen, ob das Gewicht der Person im Normalbereich liegt oder ob die Person unter- oder übergewichtig ist. Erst wenn Sie das Gewicht in Bezug zu einer Vergleichsgruppe setzen, lässt es sich interpretieren. Anders als für das Körpergewicht haben Menschen für Rohwerte psychologischer Tests üblicherweise keinen Bezugsrahmen.

Hierbei wird auch die Abhängigkeit der Interpretation von der Vergleichsgruppe deutlich. Während eine 65 kg schwere erwachsene Frau durchaus normalgewichtig sein könnte, könnte ein erwachsener Mann desselben Gewichts bereits untergewichtig und ein Kind desselben Gewichts bereits übergewichtig sein. Das Geschlecht einer Person ist nur ein möglicher, aber relativ gängiger Bezugsrahmen. Das Beispiel zeigt bereits, dass neben dem Geschlecht beispielsweise auch Körpergröße oder Alter ein Bezugsrahmen sein können.

Normskalen

Die Testergebnisse einer Normstichprobe werden anhand von Normskalen abgebildet. Normskalen orientieren sich an den Intervallen der Normalverteilung. Wesentliche Merkmale einer Normalverteilung sind Mittelwert und Standardabweichung. Die Standardnormalverteilung weist einen Mittelwert von 0 und eine Standardabweichung von 1 auf. Bei Normalverteilungen liegen:

- ◆ 68.3% aller Werte im Intervall von plus/minus einer Standardabweichung um den Mittelwert, bei der Standardnormalverteilung also im Bereich von -1 bis 1.
- ◆ 95.4% aller Werte im Intervall von plus/minus zwei Standardabweichung um den Mittelwert, bei der Standardnormalverteilung also im Bereich von -2 bis 2.

Durch verschieden breite Intervalle wird jedem möglichen Wert einer Normskala eine bestimmte Häufigkeit an Rohwerten in der Normstichprobe zugeordnet. Die Werte einer Normskala werden immer ganzzahlig, also ohne Nachkommastelle, angegeben.

T-Werte

Eine gängige Normskala ist die *T*-Skala mit einem Mittelwert von 50 und einer Standardabweichung von 10. Als gängiger Durchschnittsbereich wird häufig der Bereich von plus/minus einer Standardabweichung um den Mittelwert angegeben. Somit gelten *T*-Werte im Bereich von 40 bis 60 ($M \pm 1 s = 50 \pm 10$) als durchschnittlich, *T*-Werte kleiner als 40 ($M - 1 s = 50 - 10$) als unterdurchschnittlich und *T*-Werte größer als 60 ($M + 1 s = 50 + 10$) als überdurchschnittlich.

Interpretation

Im Durchschnittsbereich von plus/minus einer Standardabweichung um den Mittelwert befinden sich auf der *T*-Skala ca. 71% der entsprechenden Vergleichsgruppe. Auf den Unter- und Überdurchschnittsbereich der *T*-Skala entfallen die verbleibenden ca. 29% der entsprechenden Vergleichsgruppe zu gleichen Teilen, also ca. 15% auf den Unter- und ca. 15% auf den Überdurchschnittsbereich.

Prozentränge

Bei der Interpretation von Ergebnissen psychologischer Tests wird häufig auch auf Prozentränge zurückgegriffen. Jedem Wert einer Normskala kann ein bestimmter Prozentrang zugeordnet werden. Die zu einer bestimmten Normskala zugehörigen Prozentränge geben unmittelbar an, wieviel Prozent der Vergleichsgruppe im Vergleich zur Testperson gleiche oder niedrigere Werte haben und mittelbar wieviel Prozent der Vergleichsgruppe höhere Werte haben. Mit einem *T*-Wert von 50 hat eine Testperson einen Prozentrang von 52. Das bedeutet, dass 52% der Vergleichsgruppe im Vergleich zur Testperson gleiche oder niedrigere Werte und 48% höhere Werte haben.

Bei anormal verteilten Rohwerten bzw. geringer Variationsbreite der Rohwerte ist es möglich, den Rohwerten in der Normstichprobe auch nur Prozentränge zuzuordnen. In Anlehnung an den gängigen Durchschnittsbereich der *T*-Skala können dann Prozentränge bis 15 als unterdurchschnittlich und Prozentränge ab 85 als überdurchschnittlich gelten.

Der Vorteil von Prozenträngen ist ihre Anschaulichkeit, weshalb sie häufig auch bei der Rückmeldung der Ergebnisse an eine Testperson zum Einsatz kommen. Doch im Gegensatz zu Werten einer Normskala sind Prozentränge nicht metrisch, also die Abstände zwischen zwei Werten nicht gleich groß, sodass Abstände zwischen Prozenträngen auch nicht verglichen werden können. Im Durchschnittsbereich sind Prozentränge gestaucht, im unter- oder überdurchschnittlichen Bereich dagegen gestreckt. Ein Unterschied von vier Einheiten im Prozentrang entspricht im Durchschnittsbereich einem Unterschied von einer Einheit auf der *T*-Skala (*T*-Wert 50 mit PR 52 zu *T*-Wert 51 mit PR 56), im Überdurchschnittsbereich dagegen einem Unterschied von neun Einheiten (*T*-Wert 67 mit PR 96 zu *T*-Wert 76 mit PR 100). Mit dieser Besonderheit von Prozenträngen sollte bei deren Interpretation keinesfalls nachlässig umgegangen werden.

2.4.3.2. Messfehler und Reliabilität

Jedes Mal, wenn etwas gemessen wird, kann es zu Messfehlern kommen. Ein Großteil aller Messungen ist mit Messfehlern behaftet. Weichen die Ergebnisse von wiederholten Messungen des gleichen Gegenstands stärker voneinander ab, wird möglicherweise mit einem großen Messfehler gemessen. Weichen sie geringer voneinander ab, ist der Messfehler möglicherweise klein. Sind derartige Schwankungen erst ab einer hohen Nachkommastelle

beobachtbar, fallen sie womöglich überhaupt nicht auf und sind vermutlich auch vernachlässigbar. Messfehler bzw. derartige Schwankungen entstehen durch unzureichende Genauigkeit der Messung, aber auch durch unzureichende Stabilität des gemessenen Gegenstands, also einer Veränderung des Gegenstands in der Zeit, die zwischen den wiederholten Messungen vergangen ist. Je kleiner der Messfehler ausfällt, desto größer ist die Messgenauigkeit, die auch als Reliabilität bezeichnet wird (siehe Abschnitt *Reliabilität*, S. 94).

Analogie zur Veranschaulichung

Steigt eine Person unmittelbar hintereinander mehrmals auf eine Waage, ist zu erwarten, dass eine gute Waage, die auf eine Nachkommastelle genau misst, jedes Mal das gleiche Körpergewicht in Kilogramm anzeigen wird. Ist dies der Fall, misst die Waage genau und damit praktisch ohne Messfehler. Würde das angezeigte Gewicht allerdings jedes Mal in einem Bereich von 0.1 kg mehr oder weniger schwanken, misst die Waage mit unzureichender Genauigkeit. Die Waage hat einen Messfehler von ± 0.1 kg, denn es ist sehr unwahrscheinlich, dass das wahre Gewicht einer Person bei mehrmaligem Wiegen unmittelbar hintereinander um ± 0.1 kg schwankt. Schwankt das gemessene Körpergewicht einer Person, wenn sie hingegen mehrmals über einen Tag verteilt auf eine Waage steigt, kann die Ursache hierfür sowohl in einer unzureichenden Messgenauigkeit der Waage als auch in einer unzureichenden Stabilität des Körpergewichts, also einer Veränderung des wahren Körpergewichts im Tagesverlauf, liegen.

2.4.3.3. Standardmessfehler

Der Messfehler wird in Form des Standardmessfehlers angegeben (vgl. Bühner, 2010; Lienert & Raatz, 1998; Rentzsch & Schütz, 2009). Der Standardmessfehler beziffert den Messfehler in Einheiten der Standardabweichung der Normskala. Neben der Standardabweichung der Normskala fließt die Reliabilität einer Skala in die Berechnung des Standardmessfehlers ein.

Bei Fragestellungen, die den aktuellen Status der Testperson betreffen, kann die interne Konsistenz als Reliabilitätsschätzer herangezogen werden, für Prognosen sollte hingegen die Stabilität als Reliabilitätsschätzer verwendet werden (siehe Abschnitt *Reliabilität*, S. 94).

Formel zur Berechnung des Standardmessfehlers

(vgl. Bühner, 2010; Lienert & Raatz, 1998; Rentzsch & Schütz, 2009)

$$\sigma_{E_X} = \sigma_X \cdot \sqrt{1 - \rho_{tt}}$$

mit

 σ_X = Standardabweichung der Normskala ρ_{tt} = Reliabilität der Skala**2.4.3.4. Vertrauensbereiche**

Unter Hinzunahme des Standardmessfehlers kann für Messwerte die Unter- und Obergrenze eines Vertrauensbereichs ermittelt werden, der den wahren Wert einer Testperson mit einer zuvor festzulegenden Vertrauenswahrscheinlichkeit überdeckt.

Formel zur Berechnung der Unter- bzw. Obergrenze eines Vertrauensbereichs

(vgl. Bühner, 2010; Lienert & Raatz, 1998; Rentzsch & Schütz, 2009)

$$x - z_{1-\frac{\alpha}{2}} \cdot \sigma_{E_X} \text{ bzw. } x + z_{1-\frac{\alpha}{2}} \cdot \sigma_{E_X}$$

mit

 x = standardisierter Messwert $z_{1-\frac{\alpha}{2}}$ = z-Wert der Standardnormalverteilung bei zweiseitiger Fragestellung mitIrrtumswahrscheinlichkeit α σ_{E_X} = Standardmessfehler

Je höher die Vertrauenswahrscheinlichkeit, desto sicherer überdeckt ein Vertrauensbereich den wahren Wert. Vertrauensbereiche mit einer sehr hohen Vertrauenswahrscheinlichkeit fallen jedoch auch deutlich breiter aus, als Vertrauensbereiche mit einer niedrigeren Vertrauenswahrscheinlichkeit. Gängige Vertrauenswahrscheinlichkeiten sind beispielsweise 99% oder 95%, weniger gängig sind 90% oder 68%. Je niedriger die gewählte Vertrauenswahrscheinlichkeit ($1 - \alpha$) desto höher die Irrtumswahrscheinlichkeit (α), also die Wahrscheinlichkeit den Fehler zu begehen, mit dem Vertrauensbereich den wahren Wert einer Testperson fälschlicherweise nicht zu überdecken.

Analogie zur Veranschaulichung

Wird das Körpergewicht einer Person mehrmalig unmittelbar hintereinander oder im Tagesverlauf gemessen, kann das gemessene Körpergewicht schwanken. Diese Schwankung kann Folge einer unzureichenden Messgenauigkeit der benutzten Waage bzw. einer unzureichenden Stabilität des wahren Körpergewichts sein. Soll das Körpergewicht einer Person also möglichst zutreffend angegeben werden, ist es besser, anstelle des bei der letzten Messung festgestellten exakten Körpergewichts einen Bereich anzugeben, der das übliche Körpergewicht möglichst gut überdeckt.

Die Werte zur Bestimmung der *T*-Wert-Unter- und Obergrenzen von Vertrauensbereichen für sämtliche Skalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung mit Vertrauenswahrscheinlichkeiten von 90%, 95% und 99% bei einseitigen und zweiseitigen Fragestellungen finden sich in Tabelle A9 des Anhangs (siehe S. 150). Die für die Berechnung der Unter- und Obergrenzen der jeweiligen Vertrauensbereiche herangezogenen Reliabilitäten der entsprechenden Skalen (Cronbachs Alpha Koeffizienten in der Gesamtstichprobe, $N = 306$) finden sich ebenfalls in dieser Tabelle.

Berechnungsbeispiel

zur Bestimmung des 95%-Vertrauensbereichs bei zweiseitiger Fragestellung für einen *T*-Wert mit Hilfe von Tabelle A9 des Anhangs (S. 150):

Eine Testperson hat auf der Globalskala Sozial-emotionale Kompetenz des Fragebogens zur Selbsteinschätzung (SEC-SR) einen *T*-Wert von 51. Der zugehörige Wert zur Bestimmung der Unter- und Obergrenze des 95%-Vertrauensbereichs bei zweiseitiger Fragestellung von ± 5 kann Tabelle A9 des Anhangs (siehe S. 150) entnommen werden. Folglich überdeckt der 95%-Vertrauensbereich [46, 56] mit der *T*-Wert-Untergrenze von 46 ($51-5$) und der *T*-Wert-Obergrenze von 56 ($51+5$) mit einer Wahrscheinlichkeit von 95% den wahren Wert der Testperson.

Ob es sich um eine einseitige oder zweiseitige Fragestellung handelt und welche Vertrauenswahrscheinlichkeit der Fragestellung angemessen ist, liegt im Ermessen der interpretierenden Person und sollte von Nutzen und Risiken der zu treffenden Entscheidung abhängig gemacht werden. Da der Bewertung von Nutzen und Risiken verschiedener Entscheidungen sowie der damit in Verbindung stehenden Wahl von angemessenen Fragestellungen und Vertrauenswahrscheinlichkeiten kein allgemeingültiges Schema zugrunde gelegt werden kann, sollte die Interpretation möglichst auch nur von Personen mit fundierten Fachkenntnissen und Erfahrung in psychologischer Diagnostik vorgenommen werden.

2.4.4. Transformation von Rohwerten in standardisierte Werte

Der Rohwert auf einer Skala kann mit Hilfe der Normtabellen des Anhangs in einen T -Wert transformiert werden. In den Normtabellen sind auch die den T -Werten zugehörigen Prozentränge angegeben. Bei T -Werten und Prozenträngen handelt es sich, im Gegensatz zu Rohwerten, um sinnvoll interpretierbare Normwerte (siehe Abschnitt *Standardisierte Werte*, S. 53).

Da nicht alle theoretisch realisierbaren Skalenrohwerte in der Gesamtstichprobe vorhanden waren, befinden sich Sprünge in der Reihe der Rohwerte. Sprünge in der Verteilung der Rohwerte sind in den Normtabellen durch eckige Klammern kenntlich gemacht (siehe beispielsweise Tabelle A1, S. 136): geht einem Rohwert kein unmittelbar niedrigerer Rohwert voraus, findet sich vor dem Rohwert eine öffnende eckige Klammer, schließt an einen Rohwert kein unmittelbar höherer Rohwert an, findet sich nach dem Rohwert eine schließende eckige Klammer. Fehlende Rohwerte sind also ausgeklammert. Ist der erreichte Rohwert einer Testperson nicht in der entsprechenden Normtabelle enthalten, wird immer der nächste niedrigere Rohwert aus der Reihe der Rohwerte herangezogen.

Berechnungsbeispiele

zur Bestimmung des Normwerts einer Testperson mit Hilfe von Tabelle A1 des Anhangs (S. 136):

Hat eine Testperson auf der Skala Wahrnehmung eigener Emotionen des Fragebogens zur Selbsteinschätzung (PES-SR) einen Rohwert von 20 erreicht, kann der zugehörige T -Wert Tabelle A1 des Anhangs (siehe S. 136) entnommen werden, indem in der Spalte PES-SR die Zeile gesucht wird, die den Rohwert 20 enthält und dann in derselben Zeile in der ersten oder vorletzten Spalte der T -Wert von 48 abgelesen wird. Der zugehörige Prozentrang von 44 kann der letzten Spalte derselben Zeile entnommen werden.

Hat eine Testperson auf der Skala Wahrnehmung der Emotionen anderer des Fragebogens zur Selbsteinschätzung (PEA-SR) einen Rohwert von 8 erreicht, ist dieser Rohwert in der Spalte PEA-SR in Tabelle A1 des Anhangs (siehe S. 136) nicht zu finden. Die Rohwerte springen an dieser Stelle, der gesuchte Rohwert ist ausgeklammert. Daher wird der nächste niedrigere Rohwert aus der Reihe der Rohwerte der Spalte PEA-SR herangezogen. Für den Rohwert von 7 kann in der ersten oder vorletzten Spalte der Zeile der T -Wert von 24 abgelesen werden. Der zugehörige Prozentrang von 1 kann der letzten Spalte derselben Zeile entnommen werden.

Für den Fragebogen zur Selbsteinschätzung (SEC-I-SR) finden sich die Normwerte:

- ◆ für die Gesamtichprobe in Tabelle A1 des Anhangs (siehe S. 136),
- ◆ für die Substichprobe weiblichen Geschlechts in Tabelle A2 des Anhangs (siehe S. 138),
- ◆ für die Substichprobe männlichen Geschlechts in Tabelle A3 des Anhangs (siehe S. 140).

Für den Fragebogen zur Fremdeinschätzung (SEC-I-OR) finden sich die Normwerte:

- ◆ für die Gesamtichprobe in Tabelle A4 des Anhangs (siehe S. 142),
- ◆ für die Substichprobe weiblichen Geschlechts in Tabelle A5 des Anhangs (siehe S. 144),
- ◆ für die Substichprobe männlichen Geschlechts in Tabelle A6 des Anhangs (siehe S. 146).

Für den Situational Judgement Test (SEC-SJT) finden sich die Normwerte für die Gesamtstichprobe sowie für die Substichproben weiblichen und männlichen Geschlechts in Tabelle A7 des Anhangs (siehe S. 148).

Bei der Auswertung von Hand werden zunächst die jeweiligen Rohwerte in die Ergebnistabelle des Profilbogens übertragen. Anschließend werden, je nach gewählter Norm, die *T*-Werte und Prozentränge den entsprechenden Tabellen des Anhangs entnommen und ebenfalls in die Ergebnistabelle des Profilbogens eingetragen (siehe Abschnitt *Interpretation mit dem zugehörigen Profilbogen*, S. 76). Bei der computerbasierten Auswertung (siehe Abschnitt *Computerbasierte Auswertung*, S. 50) kann die Norm ausgewählt werden, die *T*-Werte und Prozentränge werden dann automatisch in die Ergebnistabelle des Profilbogens der Auswertungsdatei eingetragen.

2.4.5. Interpretation einzelner Skalen

Die Interpretation einzelner Skalen sollte primär anhand von *T*-Werten erfolgen (siehe Abschnitt *Normskalen*, S. 54).

Dabei gelten *T*-Werte:

- ◆ kleiner 40 als unterdurchschnittliche Werte,
- ◆ innerhalb des Bereichs von 40 bis 60 als durchschnittliche Werte und
- ◆ größer als 60 als überdurchschnittliche Werte.

T-Werte außerhalb des Durchschnittsbereichs können als auffällig interpretiert werden. Dabei weisen *T*-Werte kleiner als 40 auf Defizite, *T*-Werte größer als 60 auf Ressourcen hin.

Interpretation

Stark vereinfacht gilt also: je kleiner bzw. größer ein T -Wert, desto auffälliger ist dieser T -Wert.

Da die T -Werte nicht frei von Messfehlern sind (siehe Abschnitt *Messfehler und Reliabilität*, S. 55), sollte für die Interpretation ein angemessener Vertrauensbereich hinzugezogen werden (siehe Abschnitt *Vertrauensbereiche*, S. 57).

Überdeckt der Vertrauensbereich eines unterdurchschnittlichen T -Werts:

- ◆ ausschließlich den Unterdurchschnittsbereich (Obergrenze des Vertrauensbereichs < 40), kann der T -Wert als unterdurchschnittlich interpretiert werden,
- ◆ den Unterdurchschnitts- und den Durchschnittsbereich, kann der T -Wert als unterdurchschnittlich bis durchschnittlich interpretiert werden.

Überdeckt der Vertrauensbereich den T -Wert von 50, kann der T -Wert als durchschnittlich interpretiert. Überdeckt der Vertrauensbereich eines überdurchschnittlichen T -Werts:

- ◆ den Durchschnitts- und den Überdurchschnittsbereich, kann der T -Wert als durchschnittlich bis überdurchschnittlich interpretiert werden,
- ◆ ausschließlich den Überdurchschnittsbereich (Untergrenze des Vertrauensbereichs > 60), kann der T -Wert als überdurchschnittlich interpretiert werden.

Wie bereits angesprochen kann jedem Normwert und somit auch jedem T -Wert ein entsprechender Prozentrang zugeordnet werden. Der Vorteil von Prozenträngen liegt in ihrer Anschaulichkeit, doch dabei darf nicht vergessen werden, dass Prozentränge nicht metrisch sind (siehe Abschnitt *Standardisierte Werte*, S. 53).

2.4.6. Interpretation einer Profillinie

Wie bereits angesprochen, sollte in erster Linie die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR), die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) und die Globalskala des Situational Judgement Tests (SEC-SJT) interpretiert werden (siehe Abschnitt *Allgemeine Hinweise zur Interpretation*, S. 51). Durchschnittliche T -Werte auf der Globalskala SEC-SR oder SEC-OR können allerdings auf unterschiedliche Weise zustande gekommen sein.

Hat eine Testperson auf der Globalskala SEC-SR bzw. SEC-OR einen unterdurchschnittlichen oder überdurchschnittlichen T -Wert, könnte sie womöglich auf einer der Subskalen einen diesem Wert entgegengesetzten überdurchschnittlichen bzw. unterdurchschnittlichen T -Wert haben, also beispielsweise trotz eines selbsteingeschätzten globalen Defizits auch über Ressourcen verfügen. Hat eine Testperson auf der Globalskala SEC-SR bzw. SEC-OR hingegen einen durchschnittlichen T -Wert, könnte sie:

- ◆ auch auf allen oder den meisten der Subskalen durchschnittliche *T*-Werte haben oder
- ◆ auf einem Teil der Subskalen unterdurchschnittliche und auf einem anderen Teil der Subskalen überdurchschnittliche *T*-Werte, also beispielsweise trotz selbsteingeschätzter globaler Unauffälligkeit auf einem Teil der Subskalen auffällig sein.

Da die Subskalen relativ hoch interkorrelieren und damit weitestgehend homogene Merkmale erfassen, ist zwar zu erwarten, dass die Werte einer Person auf der Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) bzw. auf der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) und ihre Werte auf den jeweiligen Subskalen ähnlich ausfallen, jedoch schließt diese erhöhte Wahrscheinlichkeit nicht per se aus, dass die Werte auf den Subskalen im Einzelfall divergieren können.

Die Profillinie einer Testperson liefert diesbezüglich einen schnellen Überblick und erlaubt eine detaillierte Betrachtung der Ergebnisse der Testperson. Die Profillinie für die Selbsteinschätzung und die Fremdeinschätzung kann auf dem Profilbogen in das *Profil für SEC-I* eingezeichnet werden. Mit Hilfe der Profillinie werden die *T*-Werte einer Testperson bzw. deren „Verlauf“ kompakt und anschaulich dargestellt. Auch wenn die Werte eines Profils mit einer Profillinie verbunden werden, um den „Verlauf“ des Profils zu veranschaulichen, bildet die Profillinie dennoch keinen (zeitlichen) Verlauf im eigentlichen Sinn ab, sondern die Relation der Werte zueinander.

Zur Erstellung des Profils wird das Testergebnis einer Testperson in Form der *T*-Werte für die Globalskala und Subskalen jeweils separat für das Ergebnis aus der Selbsteinschätzung und das Ergebnis aus der Fremdeinschätzung in das *Profil für SEC-I* auf dem Profilbogen übertragen. Die *T*-Werte auf den Subskalen des Fragebogens zur Selbsteinschätzung bzw. den Subskalen des Fragebogens zur Fremdeinschätzung werden dann mit einer Linie verbunden (siehe beispielsweise Abbildung 16 im Abschnitt *Interpretationsbeispiel für ein unauffälliges Testergebnis*, S. 81). Diese Linie verbindet nur die Subskalen miteinander, die Globalskala wird nicht mit dieser Linie verbunden. Der „Verlauf“ der Profillinie veranschaulicht die Relation der Subskalen zueinander, sodass diese Relation besser interpretiert werden kann. Ebenso wie bei der Interpretation einzelner Skalen (siehe Abschnitt *Interpretation einzelner Skalen*, S. 60) sollte hierbei berücksichtigt werden, dass die *T*-Werte, auf denen die Profillinie basiert, nicht frei von Messfehlern sind (siehe Abschnitt *Messfehler und Reliabilität*, S. 55). Dementsprechend sollten auch Vertrauensbereiche für die einzelnen *T*-Werte in das *Profil für SEC-I* eingezeichnet werden (siehe Abschnitt *Vertrauensbereiche*, S. 57).

Die Profillinie der Selbsteinschätzung kann so auch in Relation zur Profillinie der Fremdeinschätzung betrachtet werden, sodass Differenzen zwischen Selbst- und Fremdeinschätzung veranschaulicht und besser interpretiert werden können (siehe beispielsweise Abbildung 16 im Abschnitt *Interpretationsbeispiel für ein unauffälliges Testergebnis*, S. 81).

Auch bei der Interpretation von Differenzen sollte berücksichtigt werden, dass die in die Differenz einfließenden T -Werte nicht frei von Messfehlern sind und somit auch die Differenz zwischen zwei T -Werten messfehlerbehaftet ist (siehe Abschnitt *Interpretation von Differenzen*, S. 63).

2.4.7. Interpretation von Differenzen

Die Inhalte dieses Abschnitts setzen Kenntnisse fachlicher Grundlagen der Inferenzstatistik voraus (unter anderem zur statistischen Bedeutsamkeit empirischer Ergebnisse, dem Fehler 1. Art, zwei- bzw. einseitigen Fragestellungen). Anders als im Abschnitt zu den Grundlagen der Interpretation psychologischer Testergebnisse (siehe S. 53) kann im Rahmen dieses Testmanuals nicht näher auf diese Grundlagen eingegangen werden. Wir empfehlen bei Unsicherheiten auf einschlägige Fachliteratur zurückzugreifen.

Sollen T -Werte unmittelbar miteinander verglichen werden, sollte dabei berücksichtigt werden, dass die T -Werte selbst und damit auch beliebige Differenzen zwischen zwei T -Werten nicht frei von Messfehlern sind (siehe Abschnitt *Messfehler und Reliabilität*, S. 55). Eine besonders konservative Möglichkeit des Vergleichs zweier T -Werte ist daher der Vergleich anhand ihrer Vertrauensbereiche: Überschneiden sich zwei Vertrauensbereiche nicht, ist mit entsprechender Vertrauenswahrscheinlichkeit auch davon auszugehen, dass der Vertrauensbereich der Differenz nicht den Wert 0 überdeckt und damit die Differenz der zwei T -Werte statistisch bedeutsam ist. Überschneiden sie sich hingegen, ist es wahrscheinlich, dass der Vertrauensbereich der Differenz auch den Wert 0 überdeckt und die vermeintliche Differenz eine Folge von Messfehlern ist. Eine weniger konservative Möglichkeit ist der Einbezug der kritischen Differenz: Überschreitet die Differenz zweier T -Werte die kritische Differenz, ist mit festgelegter Vertrauenswahrscheinlichkeit davon auszugehen, dass diese Differenz statistisch bedeutsam ist. Wird die kritische Differenz hingegen nicht überschritten bzw. ist die Differenz zweier T -Werte kleiner/gleich der kritischen Differenz, ist es wahrscheinlich, dass die vermeintliche Differenz eine Folge von Messfehlern ist.

Die kritische Differenz beziffert den messfehlerbedingten Anteil der Streuung der Differenz in Einheiten der Standardabweichung der Normskala. In die Berechnung fließen der kritische z -Wert, die Standardabweichung der Normskala und die Reliabilitäten der betrachteten Skalen ein. Bei Fragestellungen, die den aktuellen Status der Testperson betreffen, kann die interne Konsistenz als Reliabilitätsschätzer herangezogen werden, für Prognosen sollte hingegen die Stabilität als Reliabilitätsschätzer verwendet werden (siehe Abschnitt *Reliabilität*, S. 94).

Formel zur Berechnung einer kritischen Differenz

(vgl. Bühner, 2010; Lienert & Raatz, 1998; Rentzsch & Schütz, 2009)

bei Skalen mit gleicher Reliabilität

$$\Delta_{krit} = z_{1-\frac{\alpha}{2}} \cdot \sigma_X \cdot \sqrt{2 \cdot (1 - \rho_{tt})}$$

bei Skalen mit verschiedener Reliabilität

$$\Delta_{krit} = z_{1-\frac{\alpha}{2}} \cdot \sigma_X \cdot \sqrt{(1 - \rho_{tt1}) + (1 - \rho_{tt2})}$$

mit

 $z_{1-\frac{\alpha}{2}}$ = z-Wert der Standardnormalverteilung bei zweiseitiger Fragestellung mitIrrtumswahrscheinlichkeit α σ_X = Standardabweichung der Normskala ρ_{tt} = Reliabilität der Skala ρ_{tt1} bzw. ρ_{tt2} = Reliabilität der ersten bzw. zweiten Skala

Verschiedene Fragestellungen erfordern den Vergleich verschiedener *T*-Werte. Für den Fragebogen zur Selbsteinschätzung und den Fragebogen zur Fremdeinschätzung sind kritische Differenzen für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung für die Globalskalen und die Subskalen in Tabelle A10 des Anhangs (siehe S. 151) angegeben. Weitere kritische Differenzen sind für eine Vertrauenswahrscheinlichkeit von 95% bei einseitiger Fragestellung in Tabelle A11 des Anhangs (siehe S. 152), eine Vertrauenswahrscheinlichkeit von 99% bei zweiseitiger Fragestellung in Tabelle A12 des Anhangs (siehe S. 153) und eine Vertrauenswahrscheinlichkeit von 99% bei einseitiger Fragestellung in Tabelle A13 des Anhangs (siehe S. 154) angegeben.

Für Unterschiede zwischen einer Globalskala und den jeweiligen Subskalen ist die kritische Differenz fehlerbehaftet, da die Globalskala aus den jeweiligen Subskalen zusammengesetzt und die Messfehler damit nicht mehr unabhängig sind. Daher werden für Unterschiede zwischen einer Globalskala und den jeweiligen Subskalen keine kritischen

Differenzen angegeben. Auch die Profillinie für die Selbsteinschätzung oder Fremdeinschätzung wird deshalb nicht mit dem Wert auf der Globalskala verbunden (siehe Abschnitt *Interpretation einer Profillinie*, S. 61).

Theoretisch können verschiedene Differenzen berechnet werden:

- ◆ intraindividuelle Differenzen innerhalb einer Skala,
- ◆ intraindividuelle Differenzen zwischen zwei Skalen,
- ◆ interindividuelle Differenzen innerhalb einer Skala und
- ◆ interindividuelle Differenzen zwischen zwei Skalen.

Die zwar theoretisch berechenbaren interindividuellen Differenzen zwischen zwei Skalen sind jedoch praktisch gesehen als diagnostische Fragestellung inhaltlich bedeutungslos. Mögliche verschiedene Fragestellungen werden anhand der zu betrachtenden Differenzen im Folgenden strukturiert.

2.4.7.1. Intraindividuelle Differenzen innerhalb einer Skala

Intraindividuelle Differenzen innerhalb einer Skala betreffen Unterschiede bzw. Veränderungen in den *T*-Werten einer einzelnen Person auf einer Skala. Im Grundsatz beziehen sich zugehörige Fragestellungen darauf, ob sich der *T*-Wert von Person X auf Skala A vom *T*-Wert von Person X auf Skala A, der zu einem früheren Zeitpunkt erhoben wurde, unterscheidet.

Mögliche Fragestellungen sind:

- ◆ Schätzt eine Testperson ihre sozial-emotionale Kompetenz (Globalskala SEC-SR) heute anders ein als in der Vergangenheit?
- ◆ Schätzt eine Testperson ihre sozial-emotionale Kompetenz (Globalskala SEC-SR) nach einer Interventionsmaßnahme anders ein?

Da diese Fragestellungen Retrodiktion bzw. Prognose betreffen, sollte für die Berechnung der kritischen Differenz die Stabilität als Reliabilitätsschätzer herangezogen werden (siehe Abschnitt *Reliabilität*, S. 94). Da die Stabilität der von Inventar und Situational Judgement Test erfassten Merkmale jedoch nicht überprüft wurde, können über die statistische Bedeutsamkeit intraindividueller Differenzen innerhalb von Skalen keine belastbaren Aussagen getroffen werden. Allerdings liefern Konvergenzen zwischen den Skalen des Fragebogens zur Selbsteinschätzung und den Skalen des Fragebogens zur Fremdeinschätzung zumindest Anhaltspunkte auf eine bedingte Stabilität der mit dem Inventar erfassten Merkmale. Daher kann für diese Fragestellungen für die Berechnung der kritischen Differenz näherungsweise die interne Konsistenz als Reliabilitätsschätzer herangezogen werden

(siehe Abschnitt *Reliabilität*, S. 94). Die Berechnung erfolgt dann analog zu den Ausführungen des folgenden Abschnitts (siehe Abschnitt *Intraindividuelle Differenzen zwischen zwei Skalen*, S. 66). Diese Limitation sollte jedoch bei der Interpretation berücksichtigt werden.

2.4.7.2. Intraindividuelle Differenzen zwischen zwei Skalen

Intraindividuelle Differenzen zwischen zwei Skalen betreffen Unterschiede in den *T*-Werten einer einzelnen Person auf zwei verschiedenen Skalen. Im Grundsatz beziehen sich zugehörige Fragestellungen darauf, ob sich der *T*-Wert von Person X auf Skala A vom *T*-Wert von Person X auf Skala B unterscheidet.

Mögliche Fragestellungen sind:

- ◆ Schätzt eine Testperson ihre Kompetenz in der Wahrnehmung eigener Emotionen (Subskala PES-SR) anders ein als ihre Kompetenz in der Regulation eigener Emotionen (Subskala RES-SR)?
- ◆ Schätzt eine Testperson ihre Kompetenz in der Regulation eigener Emotionen (Subskala RES-SR) generell als persönliches Defizit oder persönliche Ressource ein?

Da diese Fragestellungen den aktuellen Status betreffen, kann für die Berechnung der kritischen Differenz die interne Konsistenz als Reliabilitätsschätzer herangezogen werden (siehe Abschnitt *Reliabilität*, S. 94). Für diese Fragestellungen finden sich kritische Differenzen für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung in Tabelle A10 des Anhangs (siehe S. 151), eine Vertrauenswahrscheinlichkeit von 95% bei einseitiger Fragestellung in Tabelle A11 des Anhangs (siehe S. 152), eine Vertrauenswahrscheinlichkeit von 99% bei zweiseitiger Fragestellung in Tabelle A12 des Anhangs (siehe S. 153) und eine Vertrauenswahrscheinlichkeit von 99% bei einseitiger Fragestellung in Tabelle A13 des Anhangs (siehe S. 154) jeweils im Quadranten oben links.

Berechnungsbeispiel

zur Bestimmung der statistischen Bedeutsamkeit einer intraindividuellen Differenz zwischen zwei Skalen für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung mit Hilfe der Tabelle A10 des Anhangs (S. 151):

Eine Testperson hat auf der Subskala Empathisches Verhalten des Fragebogens zur Selbsteinschätzung (EBE-SR) einen T -Wert von 56 und auf der Subskala Kognitive Empathie des Fragebogens zur Selbsteinschätzung (ECO-SR) einen T -Wert von 49. Die zugehörige kritische Differenz zur Bestimmung der statistischen Bedeutsamkeit einer intraindividuellen Differenz für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung kann Tabelle A10 des Anhangs (siehe S. 151) entnommen werden (Wert von 13). Der Betrag der Differenz der beiden T -Werte in Höhe von $|56-49| = 7$ fällt kleiner/gleich aus als die kritische Differenz der beiden Subskalen in Höhe von 13. Folglich schätzt sich die Testperson mit einer Vertrauenswahrscheinlichkeit von 95% in ihrem empathischen Verhalten (Subskala EBE-SR) nicht statistisch bedeutsam anders ein als hinsichtlich ihrer kognitiven Empathie (Subskala ECO-SR).

2.4.7.3. Interindividuelle Differenzen innerhalb einer Skala

Interindividuelle Differenzen innerhalb einer Skala betreffen Unterschiede in den T -Werten von zwei Personen auf einer Skala. Im Grundsatz beziehen sich zugehörige Fragestellungen darauf, ob sich der T -Wert von Person X auf Skala A vom T -Wert von Person Y auf Skala A unterscheidet.

Mögliche Fragestellungen sind:

- ◆ Schätzen zwei Testpersonen ihre sozial-emotionale Kompetenz (Globalskala SEC-SR) jeweils unterschiedlich ein (Vergleich von zwei Selbsteinschätzungen)?
- ◆ Schätzt eine Testperson ihre sozial-emotionale Kompetenz selbst (Globalskala SEC-SR) anders ein als diese von einer anderen Person (Globalskala SEC-OR) eingeschätzt wird (Vergleich von Selbst- und Fremdeinschätzung)?
- ◆ Schätzen zwei Personen die sozial-emotionale Kompetenz einer Testperson (Globalskala SEC-OR) unterschiedlich ein (Vergleich von zwei Fremdeinschätzungen)?

Da diese Fragestellungen ebenfalls den aktuellen Status betreffen, kann für die Berechnung der kritischen Differenz, analog zu intraindividuelle Differenzen zwischen zwei Skalen, die interne Konsistenz als Reliabilitätsschätzer herangezogen werden (siehe Abschnitt *Reliabilität*, S. 94). Für diese Fragestellungen finden sich kritische Differenzen für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung in Tabelle A10 des Anhangs

(siehe S. 151), eine Vertrauenswahrscheinlichkeit von 95% bei einseitiger Fragestellung in Tabelle A11 des Anhangs (siehe S. 152), eine Vertrauenswahrscheinlichkeit von 99% bei zweiseitiger Fragestellung in Tabelle A12 des Anhangs (siehe S. 153) und eine Vertrauenswahrscheinlichkeit von 99% bei einseitiger Fragestellung in Tabelle A13 des Anhangs (siehe S. 154):

- ◆ im Quadranten oben links für den Vergleich von zwei Selbsteinschätzungen,
- ◆ im Quadrant unten links für den Vergleich von Selbst- und Fremdeinschätzung,
- ◆ im Quadrant unten rechts für den Vergleich von zwei Fremdeinschätzungen.

Berechnungsbeispiel

zur Bestimmung der statistischen Bedeutsamkeit einer intraindividuellen Differenz innerhalb einer Skala für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung mit Hilfe der Tabelle A10 des Anhangs (S. 151):

Eine Testperson hat auf der Subskala *Positive Beziehungsgestaltung* des Fragebogens zur Selbsteinschätzung (SOB-SR) einen *T*-Wert von 44 und auf der inhaltsgleichen Skala des Fragebogens zur Fremdeinschätzung (SOB-OR) einen *T*-Wert von 57. Die zugehörige kritische Differenz zur Bestimmung der statistischen Bedeutsamkeit einer intraindividuellen Differenz für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung kann Tabelle A10 des Anhangs (siehe S. 151) entnommen werden (Wert von 12). Der Betrag der Differenz der beiden *T*-Werte in Höhe von $|44-57| = 13$ fällt größer aus als die kritische Differenz der beiden Subskalen in Höhe von 12. Folglich schätzt sich die Testperson mit einer Vertrauenswahrscheinlichkeit von 95% in ihrer positiven Beziehungsgestaltung (Subskala SOB-SR) statistisch bedeutsam schlechter ein, als sie von einer anderen Person (Subskala SOB-OR) eingeschätzt wird.

2.4.8. Inhaltliche Beschreibung der Skalen

Da der Fragebogen zur Selbsteinschätzung und der Fragebogen zur Fremdeinschätzung so konstruiert wurden, dass sie in den erfassten Merkmalen inhaltlich übereinstimmen, können die Globalskala und die Subskalen parallel interpretiert werden – jeweils als Selbst- bzw. Fremdeinschätzung der entsprechenden Fähigkeit. Dabei sollte jedoch beachtet werden, dass es sich bei Selbst- und Fremdeinschätzungen um subjektive Einschätzungen der eigenen bzw. einer anderen Person handelt und nicht um objektive Ergebnisse aus einem Leistungstest (siehe Abschnitt *Theoretischer Hintergrund*, S. 18; und Abschnitt *Allgemeine*

Hinweise zur Interpretation, S. 51). Eine Fremdeinschätzung kann jedoch durch den Einbezug zusätzlicher Fremdeinschätzungen der Testperson durch weitere Personen zunehmend objektiviert werden.

Im Gegensatz zu den Skalen des Fragebogens zur Selbsteinschätzung und denen des Fragebogens zur Fremdeinschätzung, die subjektive Einschätzungen erfassen, erfasst die Globalskala des Situational Judgement Test (SEC-SJT) eine objektive Leistung – das Potential (Wissen bzw. Erfahrung) einer Testperson, sich in verschiedensten Situationen sozial-emotional kompetent verhalten zu können. Das Vorhandensein dieses Potentials bedeutet allerdings nicht, dass dieses zwingend in den üblichen (teils automatisierten) Verhaltensgewohnheiten zum Ausdruck kommt bzw. dass die Testperson kongruent zu ihrem Wissen bzw. ihren Erfahrungen handelt (siehe Abschnitt *Theoretischer Hintergrund*, S. 18).

2.4.8.1. Skala des Situational Judgement Tests

Globalskala Sozial-emotionale Kompetenz (SEC-SJT)

Die Globalskala Sozial-emotionale Kompetenz des Situational Judgement Test (SEC-SJT) erfasst mit 10 Items Wissen bzw. Erfahrung im Umgang mit den Emotionen und Bedürfnissen anderer Personen wie auch der eigenen Person.

Testpersonen mit einem hohen Prozentrang verfügen über mehr (intuitives) Wissen bezüglich des Umgangs mit den Emotionen und Bedürfnissen anderer Personen wie auch der eigenen Person als andere.

Testpersonen mit einem niedrigen Prozentrang verfügen über weniger (intuitives) Wissen im Umgang mit den Emotionen und Bedürfnissen anderer Personen wie auch der eigenen Person als andere.

2.4.8.2. Skalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung

Globalskala Sozial-emotionale Kompetenz (SEC-SR bzw. SEC-OR)

Die Globalskala Sozial-emotionale Kompetenz des Fragebogens zur Selbsteinschätzung (SEC-SR) bzw. des Fragebogens zur Fremdeinschätzung (SEC-OR) erfasst mit 41 Items den Umgang mit den Emotionen und Bedürfnissen anderer Personen wie auch der eigenen Person, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala achten deutlich stärker als andere auf das emotionale Wohlbefinden anderer Personen. Sie pflegen einen besonders mitfühlenden, unterstützenden und wertschätzenden Umgang mit anderen Personen. Tendenziell profitiert ihr Umgang mit den Emotionen und Bedürfnissen anderer Personen davon, dass sie Emotionen differenzierter wahrnehmen können, einfühlsamer damit umgehen können und seltener von eigenen Emotionen überwältigt werden.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala achten deutlich weniger als andere auf das emotionale Wohlbefinden anderer Personen. Sie sind anderen Personen gegenüber weniger mitfühlend, unterstützend und wertschätzend. Tendenziell haben sie häufiger Schwierigkeiten darin, Emotionen differenziert wahrzunehmen oder werden häufiger von eigenen Emotionen überwältigt, was sie in ihrer Einfühlsamkeit beeinträchtigt.

Subskala Empathisches Verhalten (EBE-SR bzw. EBE-OR)

Die Subskala Empathisches Verhalten des Fragebogens zur Selbsteinschätzung (EBE-SR) bzw. des Fragebogens zur Fremdeinschätzung (EBE-OR) erfasst mit 4 Items die Bereitschaft, andere Personen aus eigenem Mitgefühl heraus zu unterstützen und ihnen Hilfe zukommen zu lassen, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala sind weitaus mehr als andere dazu bereit, andere Personen zu unterstützen und ihnen Hilfe zukommen zu lassen, wenn sie das Gefühl haben, dass die andere Person dies braucht.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala sind deutlich weniger als andere dazu bereit, anderen Personen, die emotional belastet sind, Unterstützung zukommen zu lassen.

Subskala Kognitive Empathie (ECO-SR bzw. ECO-OR)

Die Subskala Kognitive Empathie des Fragebogens zur Selbsteinschätzung (ECO-SR) bzw. des Fragebogens zur Fremdeinschätzung (ECO-OR) erfasst mit 5 Items die Fähigkeit, sich gedanklich in andere Personen hineinversetzen und vermuten zu können, was Gründe oder Auslöser ihres emotionalen Erlebens sind, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala können sich gedanklich deutlich besser als andere in andere Personen hineinversetzen und sich weitaus besser vorstellen, was Gründe oder Auslöser für die Emotionen anderer sind.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala fällt es deutlich schwerer als anderen, die Perspektive anderer Personen einzunehmen und sich vorzustellen, was Gründe oder Auslöser für deren emotionales Erleben sind.

Subskala Wahrnehmung der Emotionen anderer (PEA-SR bzw. PEA-OR)

Die Subskala Wahrnehmung der Emotionen anderer des Fragebogens zur Selbsteinschätzung (PEA-SR) bzw. des Fragebogens zur Fremdeinschätzung (PEA-OR) erfasst mit 5 Items die Fähigkeit, emotionales Erleben bei anderen Personen wahrnehmen und einordnen zu können, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala können das emotionale Erleben anderer Personen deutlich besser wahrnehmen und die Emotionen anderer deutlich besser einordnen als andere.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala fällt es wesentlich schwerer als anderen, emotionales Erleben bei anderen Personen zu erkennen und einzuordnen.

Subskala Wahrnehmung eigener Emotionen (PES-SR bzw. PES-OR)

Die Subskala Wahrnehmung eigener Emotionen des Fragebogens zur Selbsteinschätzung (PES-SR) bzw. des Fragebogens zur Fremdeinschätzung (PES-OR) erfasst mit 5 Items die Fähigkeit, eigenes emotionales Erleben wahrnehmen und einordnen zu können, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala können ihr eigenes emotionales Erleben deutlich differenzierter wahrnehmen und ihre Emotionen deutlich besser einordnen als andere.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala nehmen ihr emotionales Erleben weitaus weniger differenziert wahr, es fällt ihnen deutlich schwerer als anderen, angenehmes oder unangenehmes emotionales Erleben auf unterschiedliche Emotionen beziehen zu können.

Subskala Regulation der Emotionen anderer (REA-SR bzw. REA-OR)

Die Subskala Regulation der Emotionen anderer des Fragebogens zur Selbsteinschätzung (REA-SR) bzw. des Fragebogens zur Fremdeinschätzung (REA-OR) erfasst mit 5 Items die Fähigkeit, regulierend auf das emotionale Erleben anderer Personen einwirken und sie bei der Bewältigung unangenehmen emotionalen Erlebens unterstützen zu können, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala können deutlich besser als andere das emotionale Erleben anderer Personen beeinflussen und sie darin unterstützen, mit ihren unangenehmen Emotionen umzugehen.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala sind weitaus häufiger als andere damit überfordert, anderen Personen zu helfen, wenn diese von ihren Emotionen überwältigt werden.

Subskala Regulation eigener Emotionen (RES-SR bzw. RES-OR)

Die Subskala Regulation eigener Emotionen des Fragebogens zur Selbsteinschätzung (RES-SR) bzw. des Fragebogens zur Fremdeinschätzung (RES-OR) erfasst mit 4 Items die Fähigkeit, regulierend auf das eigene emotionale Erleben einwirken und unangenehme Emotionen bewältigen zu können, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala können deutlich besser als andere ihr emotionales Erleben selbst beeinflussen und mit unangenehmen Emotionen umgehen.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala werden wesentlich schneller als andere von ihren eigenen Emotionen überwältigt und haben größere Schwierigkeiten im Umgang mit eigenen unangenehmen Emotionen.

Subskala Positive Beziehungsgestaltung (SOB-SR bzw. SOB-OR)

Die Subskala Positive Beziehungsgestaltung des Fragebogens zur Selbsteinschätzung (SOB-SR) bzw. des Fragebogens zur Fremdeinschätzung (SOB-OR) erfasst mit 9 Items das Bemühen um den Aufbau und die Aufrechterhaltung positiver Beziehungen zu anderen Personen und die Übernahme von Verantwortung für das Wohlergehen anderer Personen, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala sind deutlich stärker als andere darum bemüht, positive Beziehungen zu anderen Personen aufzubauen und sind besorgter um das Wohlergehen anderer Personen.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala sind positive Beziehungen zu anderen Personen und das Wohlergehen anderer Personen weitaus weniger wichtig als anderen.

Subskala Beachtung sozialer Normen (SON-SR bzw. SON-OR)

Die Subskala Beachtung sozialer Normen des Fragebogens zur Selbsteinschätzung (SON-SR) bzw. des Fragebogens zur Fremdeinschätzung (SON-OR) erfasst mit 4 Items die Achtung und Anerkennung anderer Personen durch die Einhaltung gängiger sozialer Verhaltensregeln, jeweils im Selbstbild bzw. Fremdbild.

Testpersonen mit überdurchschnittlicher Ausprägung auf dieser Skala achten deutlich mehr als andere darauf, mit anderen Personen wertschätzend und respektvoll umzugehen.

Testpersonen mit unterdurchschnittlicher Ausprägung auf dieser Skala gehen mit anderen Personen deutlich weniger wertschätzend und respektloser um als andere.

2.4.9. Situational Judgement Test als Screeningverfahren

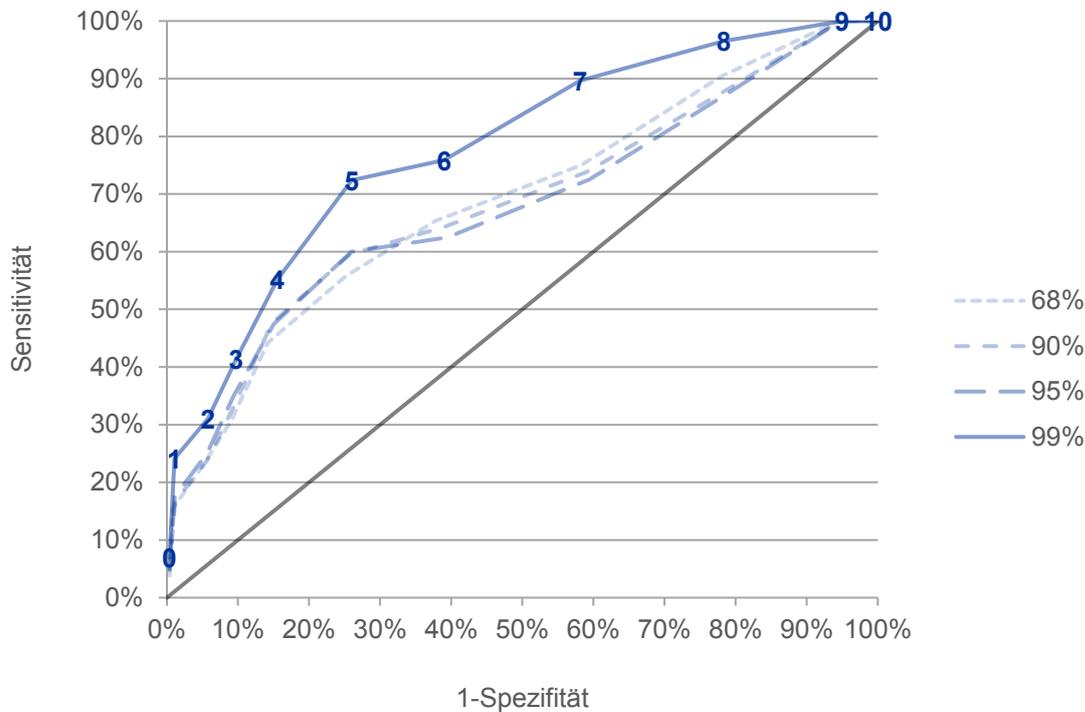
Der Situational Judgement Test kann auch als Screeningverfahren im Sinne der kriteriumsorientierten Diagnostik eingesetzt werden. Anders als in der normorientierten Diagnostik, in der Testergebnisse standardisiert werden, indem sie in Bezug zu einer Vergleichsgruppe gesetzt werden (siehe Abschnitt *Standardisierte Werte*, S. 53), wird in der kriteriumsorientierten Diagnostik in der Regel geprüft, ob Testergebnisse einen zuvor festgelegten Cut-Off-Wert überschreiten. Aufgrund der relativ hohen Itemschwierigkeiten und der damit einhergehenden linksschiefen Verteilung der Rohwerte der Globalskala des Situational Judgement Tests in der Gesamtstichprobe (siehe Abschnitt *Analyse der Items des Situational Judgement Tests*, S. 91; und Abschnitt *Analyse der Globalskala des Situational Judgement Tests*, S. 92) kann das Testergebnis vor allem Hinweise auf potentielle Defizite geben (anders als in der Umgangssprache weist ein dichotomes Item in der klassischen Testtheorie eine hohe Itemschwierigkeit auf, wenn es von einem Großteil der Testpersonen korrekt gelöst wurde). Liefert der Situational Judgement Test Hinweise auf ein potentielles Defizit, sollten der Fragebogen zur Fremdeinschätzung und wenn möglich der Fragebogen zur Selbsteinschätzung zur weiteren Abklärung genutzt werden.

Um einen geeigneten Cut-Off-Wert festlegen zu können, werden meist Receiver Operating Characteristic Kurven (ROC-Kurven) berechnet. Eine ROC-Kurve beschreibt das Verhältnis von Richtig-Positiv-Rate (Sensitivität) zu Falsch-Positiv-Rate (1-Spezifität) für alle möglichen Werte eines Klassifikators – also die Klassifikationsgüte eines diagnostischen Verfahrens bezogen auf die verschiedenen Werte des Klassifikators (Macmillan & Creelman, 2005). Für ein rein zufällig klassifizierendes diagnostisches Verfahren entspricht die theoretische ROC-Kurve der Winkelhalbierenden. Je besser ein diagnostisches Verfahren, desto größer die Fläche zwischen ROC-Kurve und Winkelhalbierender. Bei einem perfekt klassifizierenden diagnostischen Verfahren verläuft die ROC-Kurve direkt auf der y-Achse, sodass bei einer Sensitivität von 100% eine Falsch-Positiv-Rate von 0% erreicht wird. Je weiter ein Klassifikator von der Winkelhalbierenden entfernt ist, desto besser ist das Verhältnis aus Sensitivität und Falsch-Positiv-Rate für diesen Klassifikator.

Die ROC-Kurven in Abbildung 11 (auf der nächsten Seite) veranschaulichen Sensitivität und Falsch-Positiv-Rate für eine Auffälligkeit auf der Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) oder auf der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) bei verschiedenen Vertrauenswahrscheinlichkeiten.

Abbildung 11

Richtig-Positiv-Rate (Sensitivität) und Falsch-Positiv-Rate (1-Spezifität) für eine Auffälligkeit auf einer der Globalskalen der Selbst- oder Fremdeinschätzung (SEC-SR oder SEC-OR) bei verschiedenen Vertrauenswahrscheinlichkeiten



Anmerkung. $N = 306$. ROC-Kurven für Vertrauenswahrscheinlichkeiten von 68%, 90%, 95% und 99% bei zweiseitiger Fragestellung. Eine Person galt als auffällig, wenn der Vertrauensbereich für einen auffälligen (unterdurchschnittlichen) T -Wert für die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) oder die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) nicht den Durchschnittsbereich überschneidet.

Abbildung 11 kann entnommen werden, dass für einen Cut-Off-Wert von 5 (Rohwert im Situational Judgement Test, entspricht einem Prozentrang von 30 in der Gesamtstichprobe bzw. einem Prozentrang von 22 bzw. 35 in den Substichproben weiblichen bzw. männlichen Geschlechts) der beste Trade-Off zwischen Sensitivität und Falsch-Positiv-Rate erreicht werden kann. Wird der Situational Judgement Test als Screeningverfahren eingesetzt, könnte demzufolge ein Rohwert kleiner/gleich 5 ein Indikator für die zusätzliche Erhebung des Fragebogens zur Fremdeinschätzung und wenn möglich des Fragebogens zur Selbsteinschätzung sein.

Ob der Situational Judgement Test für die normorientierte oder kriteriumsorientierte Diagnostik eingesetzt werden soll, hängt ebenso wie die Wahl eines geeigneten Cut-Off-Wertes von der Fragestellung und den zur Verfügung stehenden Ressourcen (beispielsweise Personal, Zeit, finanziellen Mitteln) ab. Stehen ausreichend Ressourcen zur Verfügung kann der

Situational Judgement Test im Rahmen der normorientierten Diagnostik zusätzlich zum Fragebogen zur Selbsteinschätzung bzw. zum Fragebogen zur Fremdeinschätzung erhoben werden, sodass die Ergebnisse integriert werden können und die diagnostischen Verfahren ihr volles Potential entfalten. Stehen nur begrenzt Ressourcen zur Verfügung kann der Situational Judgement Test als Screeningverfahren vor dem Fragebogen zur Selbsteinschätzung oder dem Fragebogen zur Fremdeinschätzung erhoben werden und als Indikator für die etwaige Notwendigkeit einer zusätzlichen Erhebung einer Selbsteinschätzung bzw. einer oder mehrerer Fremdeinschätzung(en) dienen. Da derartigen Entscheidungen kein allgemeingültiges Schema zugrunde gelegt werden kann, sollten diese möglichst nur von entsprechend ausgebildetem Fachpersonal getroffen werden.

2.4.10. Ergebnisrückmeldungen

Die Anwendung von Inventar und Situational Judgement Test sollte nie dem Selbstzweck dienen, sondern der erste Schritt einer zielgerichteten Intervention sein. Schließlich wurden die diagnostischen Verfahren speziell für die gezielte ressourcenorientierte Förderung sozial-emotionaler Kompetenz(en) bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelt.

Bereits bei der Rückmeldung von Testergebnissen handelt es sich um eine Intervention, über deren Ziel sich die rückmeldende Person zuvor Gedanken machen sollte. Es sollte auch nicht unterschätzt werden, dass die Möglichkeit, eine Rückmeldung zu den Ergebnissen zu erhalten, ein Anreiz für die sorgfältige Bearbeitung sowie offenes und ehrliches Antworten im Fragebogen zur Selbsteinschätzung sein kann (siehe Abschnitt *Verzerrungen im Antwortverhalten*, S. 36).

Ergebnisrückmeldungen können der Testperson dabei helfen, sich selbst (besser) zu verstehen. Im Fall der Rückmeldung von Ergebnissen aus einer oder mehreren Fremdeinschätzungen können Ergebnisrückmeldungen auch dabei helfen, zu verstehen, wie man von anderen Personen wahrgenommen wird. Jedoch ist die Rückmeldung von Ergebnissen, insbesondere auch die Rückmeldung von Ergebnissen aus einer Fremdeinschätzung, mit gewissen Gefahren behaftet. Dennoch sollte Personen, die psychologische Tests bearbeitet haben, im Anschluss ein professionelles Beratungsgespräch mit einer psychologisch ausgebildeten Beratungsperson angeboten werden. Im Rahmen des Beratungsgesprächs können dann individuelle Testergebnisse besprochen werden, sofern die Testperson diesen Wunsch äußert. Beratungsgespräche, wie auch Rückmeldungen von Testergebnissen, sollten ausschließlich auf Wunsch der Testperson stattfinden und nicht aufgezwungen werden. Zeigt die Testperson nach der Bearbeitung des Fragebogens zur Selbsteinschätzung oder des Situational Judgement Tests Gesprächsbedarf, sollte also auf diesen Bedarf eingegangen werden.

Die rückmeldende Person sollte Nutzen und Gefahren der Rückmeldung von Testergebnissen sorgfältig abwägen und die Rückmeldung in ein konstruktives und zielgerichtetes Beratungsgespräch einbetten. Die Rückmeldung von unterdurchschnittlichen Testergebnissen, insbesondere aus der Fremdeinschätzung, sollte stets behutsam erfolgen. Für die Rückmeldung von Ergebnissen bieten sich insbesondere Prozentränge an, da diese relativ anschaulich sind (siehe Abschnitt *Standardisierte Werte*, S. 53). Beispielsweise kann gegenüber der Testperson bei einem unterdurchschnittlichen Testergebnis geäußert werden, dass sich nur wenige andere Personen ebenso niedrig einschätzen. Bei einem durchschnittlichen Testergebnis kann rückgemeldet werden, dass sich die meisten anderen Personen sehr ähnlich einschätzen. Bei einem überdurchschnittlichen Testergebnis kann geäußert werden, dass sich nur wenige andere Personen höher einschätzen.

2.4.11. Interpretation mit dem zugehörigen Profilbogen

Da ein Teil der Auswertung bereits grundlegende Entscheidungen für die Interpretation erfordert (Wahl der geeigneten Norm, Wahl der Vertrauenswahrscheinlichkeit) und damit die Auswertung fließend in die Interpretation übergeht, wird die jeweils begonnene Beschreibung der Auswertung aus dem Abschnitt *Auswertung mittels Auswertungsbögen* (siehe S. 43) in diesem Abschnitt fortgesetzt.

2.4.11.1. Interpretation des Fragebogens zur Selbsteinschätzung

In den bereits erfolgten Schritten 1 bis 5 (siehe Abschnitt *Auswertung des Fragebogens zur Selbsteinschätzung*, S. 43) wurden die numerischen Werte der Antworten aus dem Fragebogen zur Selbsteinschätzung in den *SEC-I-SR Auswertungsbögen* eingetragen, die Rohwerte der Subskalen und der Globalskala berechnet und in den Profilbogen übertragen.

Schritt 6: Ermittlung der Normwerte

Als Norm wurde die Gesamtstichprobe gewählt. Für den Rohwert 153 der Globalsskala und die Rohwerte der Subskalen (beispielsweise EBE-SR mit Rohwert 17) können die *T*-Werte und Prozentränge (beispielsweise für die Globalsskala ein *T*-Wert von 51 und ein Prozentrang von 56, für die Subskala EBE-SR ein *T*-Wert von 56 und ein Prozentrang von 74) für die Gesamtstichprobe Tabelle A1 des Anhangs (siehe S. 136) entnommen und in den Profilbogen eingetragen werden. Schritt 6 ist in Abbildung 12 (auf der nächsten Seite) dargestellt.

Damit zu einem späteren Zeitpunkt nachvollzogen werden kann, welche Norm (die Gesamtstichprobe bzw. die Substichprobe weiblichen oder männlichen Geschlechts) für die Ermittlung der Normwerte herangezogen wurde, wird die verwendete Norm auf dem

Interpretation

Profilbogen durch Ankreuzen dokumentiert (hier im Beispiel durch das Kreuz bei Gesamt). Die gewählte Norm ist dann auch für die Ermittlung der Normwerte für den Situational Judgement Test und den Fragebogen zur Fremdeinschätzung heranzuziehen.

Abbildung 12

Schritt 6 der Auswertung des Fragebogens zur Selbsteinschätzung mittels Profilbogen und Normtabelle



PROFILBOGEN

A

Tabelle A1

Normierungsschritte

- Übertragen Sie die Rohwerte (Summen aus dem SEC-I-SR, SEC-I-OR und SEC-SJT Auswertungsbogen) der jeweiligen Skalen (Subskalen und Globalskalen) in die dafür vorgesehenen Zellen der **Ergebnistabelle**.
- Wählen Sie eine geeignete Norm (Gesamtstichprobe, Stichprobe weiblichen bzw. männlichen Geschlechts) und entnehmen Sie dann die zugehörigen T-Werte bzw. Prozentränge (PR) dem Testmanual.

Ergebnistabelle

Norm: Gesamt weiblich männlich

	Rohwert	T-Wert	PR		Rohwert	T-Wert	PR
Skala				Skala			
SEC-SJT				SEC-OR			
SEC-SR	153	51	50	SEC-OR			
EBE-SR	77	56	74	EBE-OR			
ECO-SR	77	49	48	ECO-OR			
PEA-SR	18	49	48	PEA-OR			
PES-SR	21	50	52	PES-OR			
REA-SR	16	48	44	REA-OR			
RES-SR	16	57	77	RES-OR			
SOB-SR	30	44	29	SOB-OR			
SON-SR	18	50	52	SON-OR			

Tabelle A1

Normwerte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) für Jugendliche und junge Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15,0 bis 24,11 Jahren (N = 306)

T	SEC-SR	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	T PR
79	[204]							45]		79 100
78										78 100
77										77 100
76	[198]									76 100
75			25]			25]		44		75 99
74	194]			25]						74 99
73	193									73 99
72								43		72 99
71	192						20]			71 98
70	190-191					24		42		70 98
69	189		24							69 97
68	[186-188			24						68 97
67	183-184]					23				67 96
66	182	20]					19	41		66 95
65	181		23		25]	22				65 94
64	179-180			23						64 93
63	178		22				18	20]		63 91
62	175-177					21		40		62 89
61	173-174	19	21				17			61 87
60	171-172			22				39		60 85
59	169-170	18			24	20		38		59 83
58	166-168									58 80
57	165		20	23		19	16	37		57 77
56	163-164	17			23					56 74
55	161-162		19				16	19		55 71
54	159-160			20			15			54 67
53	157-158				22			35		53 64
52	155-156	16	18	19			14			52 60
51	153-154					17		34		51 56
50	151-152					21		18	50 52	
49	150	15	17	18			13	33		49 48
48	148-149				20	16				48 44
47	146-147						12	32	17	47 40

(Fortsetzung)

3 von 5



136

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

Schritt 7: Übertragen der Normwerte in das Profil

Für die Darstellung der Testergebnisse werden die T-Werte aus der Ergebnistabelle auf Seite 3 des Profilbogens in das *Profil für SEC-I* auf Seite 4 übertragen. Schritt 7 ist in Abbildung 13 (auf der nächsten Seite) dargestellt.

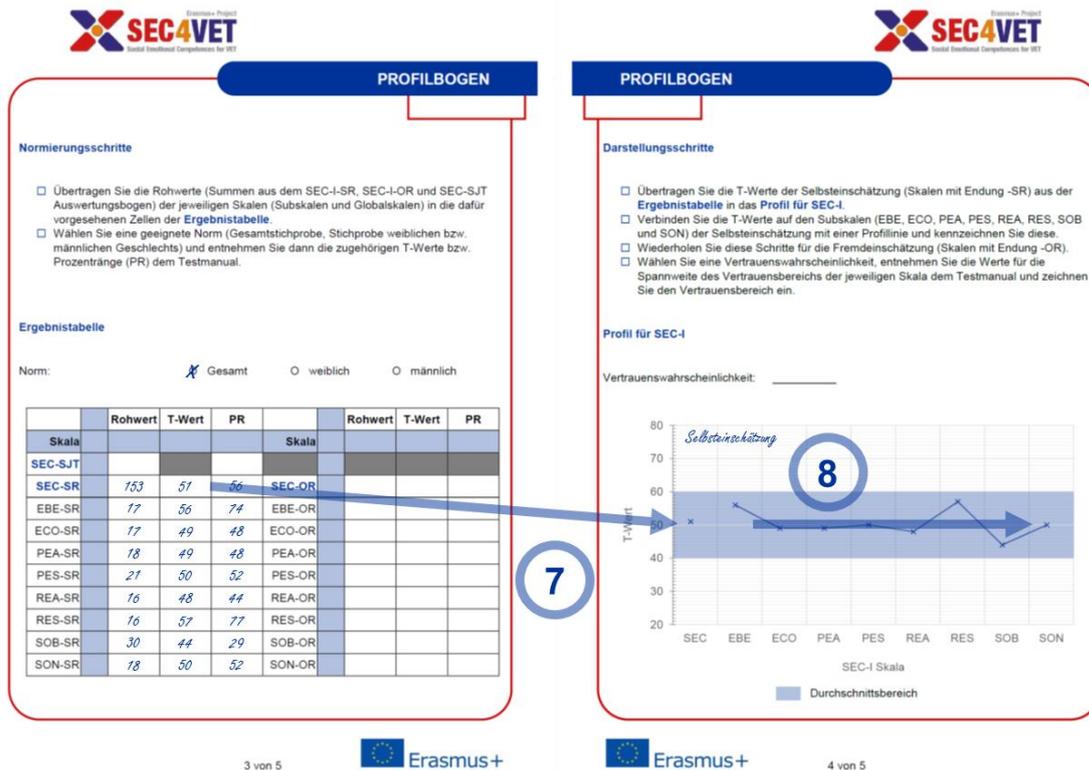
Schritt 8: Einzeichnen der Profillinie für die Normwerte

Nachdem alle T-Werte in das *Profil für SEC-I* übertragen wurden, werden die T-Werte auf den Subskalen mit einer Profillinie verbunden. Für den Fall, dass später auch Testergebnisse aus der Fremdeinschätzung dargestellt werden sollen, wird kenntlich gemacht, dass es sich um die Profillinie der Selbsteinschätzung handelt (hier im Beispiel

durch den handschriftlichen Eintrag *Selbsteinschätzung* oberhalb der Profillinie). Schritt 8 ist in Abbildung 13 dargestellt.

Abbildung 13

Schritte 7 und 8 der Auswertung des Fragebogens zur Selbsteinschätzung (SEC-I-SR) mittels Profilbogen



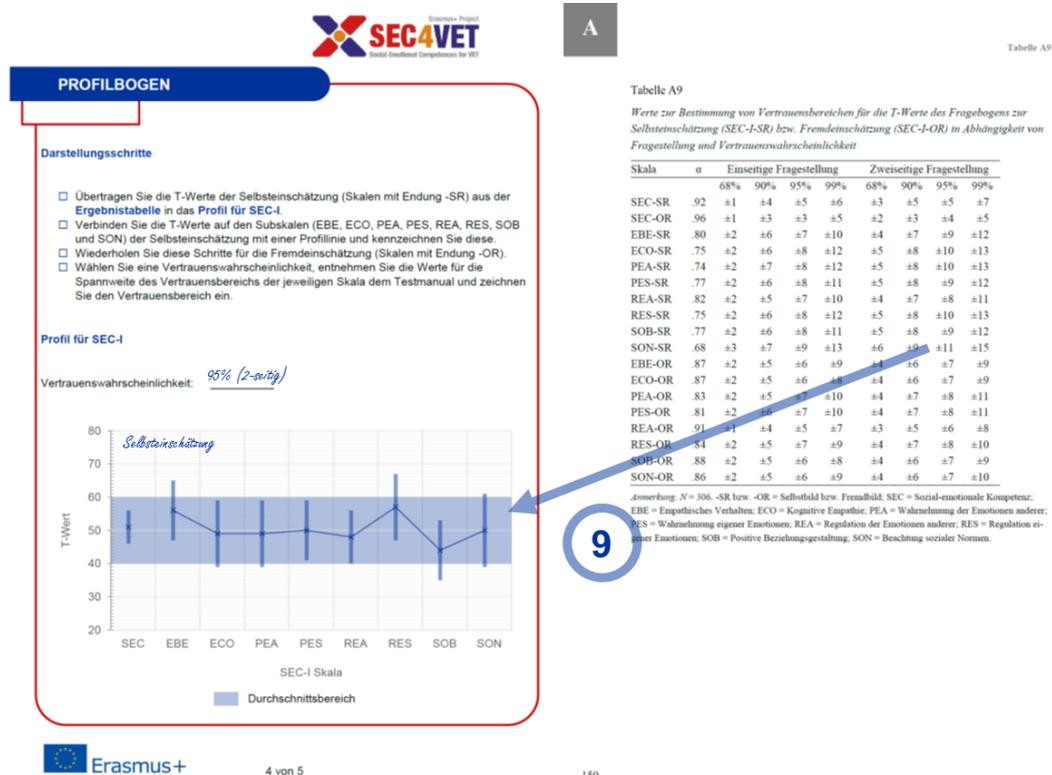
Anmerkung: Testmaterialien einer Person mit einem unauffälligen Testergebnis.

Schritt 9: Einzeichnen der Vertrauensbereiche für die Normwerte

Um die Vertrauensbereiche der Normwerte bei zweiseitiger Fragestellung mit 95% Vertrauenswahrscheinlichkeit einzeichnen zu können, werden die Werte für die Spannweite des Vertrauensbereichs der jeweiligen Skala Tabelle A9 des Anhangs (siehe S. 150) entnommen. Damit zu einem späteren Zeitpunkt nachvollzogen werden kann, welche Vertrauenswahrscheinlichkeit gewählt wurde, wird die Vertrauenswahrscheinlichkeit auf dem Profilbogen dokumentiert (hier im Beispiel durch den handschriftlichen Eintrag *95% (2-seitig)* im Feld Vertrauenswahrscheinlichkeit). Schritt 9 ist in Abbildung 14 (auf der nächsten Seite) dargestellt.

Abbildung 14

Schritt 9 der Auswertung des Fragebogens zur Selbsteinschätzung (SEC-I-SR) mittels Profilbogen und der Tabelle der Werte zur Bestimmung von Vertrauensbereichen



Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

2.4.11.2. Interpretation des Fragebogens zur Fremdeinschätzung

Da der Fragebogen zur Fremdeinschätzung parallel zum Fragebogen zur Selbsteinschätzung konstruiert wurde, erfolgt die Interpretation analog (siehe Abschnitt *Interpretation des Fragebogens zur Selbsteinschätzung*, S. 76).

2.4.11.3. Interpretation des Situational Judgement Tests

In den bereits erfolgten Schritten 1 bis 3 (siehe Abschnitt *Auswertung des Situational Judgement Test*, S. 47) wurden die numerischen Werte der Antworten aus dem Situational Judgement Test in den *SEC-SJT Auswertungsbogen* eingetragen, der Rohwert der Globalskala SEC-SJT berechnet und in den Profilbogen übertragen.

Schritt 4: Ermittlung des Normwerts

Als Norm wurde die Gesamtstichprobe gewählt. Für den Rohwert 8 auf der Globalskala SEC-SJT kann der Prozentrang von 80 für die Gesamtstichprobe Tabelle A7 des Anhangs (siehe S. 148) entnommen und in den Profilbogen eingetragen werden. Schritt 4 ist in Abbildung 15 dargestellt.

Wurden bereits andere Normwerte in die Ergebnistabelle eingetragen, ist die gewählte Norm (die Gesamtstichprobe bzw. die Substichprobe weiblichen oder männlichen Geschlechts) auch für die Ermittlung des Normwerts für die Globalskala SEC-SJT zu übernehmen. Wurden noch keine anderen Testergebnisse in die Ergebnistabelle eingetragen, wird die verwendete Norm auf dem Profilbogen durch Ankreuzen dokumentiert (hier im Beispiel durch das Kreuz bei Gesamt).

Abbildung 15

Schritt 4 der Auswertung des Situational Judgement Tests (SEC-SJT) mittels SEC-SJT Auswertungsbogen und Normtabelle

PROFILBOGEN

Normierungsschritte

- Übertragen Sie die Rohwerte (Summen aus dem SEC-I-SR, SEC-I-OR und SEC-SJT Auswertungsbogen) der jeweiligen Skalen (Subskalen und Globalskalen) in die dafür vorgesehenen Zellen der **Ergebnistabelle**.
- Wählen Sie eine geeignete Norm (Gesamtstichprobe, Stichprobe weiblichen bzw. männlichen Geschlechts) und entnehmen Sie dann die zugehörigen T-Werte bzw. Prozenträge (PR) dem Testmanual.

Ergebnistabelle

Norm: Gesamt weiblich männlich

Skala	Rohwert	T-Wert	PR	Skala	Rohwert	T-Wert	PR
SEC-SJT	8		80				
SEC-SR				SEC-OR			
EBE-SR				EBE-OR			
ECO-SR				ECO-OR			
PEA-SR				PEA-OR			
PES-SR				PES-OR			
REA-SR				REA-OR			
RES-SR				RES-OR			
SOB-SR				SOB-OR			
SON-SR				SON-OR			

Tabelle A7

Normwerte des Situational Judgement Tests (SEC-SJT) für Jugendliche und junge Erwachsenen beider Geschlechter ($N_{\text{gesamt}} = 306$), weiblichen Geschlechts ($n_{\text{weiblich}} = 111$) bzw. männlichen Geschlechts ($n_{\text{männlich}} = 195$) mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren

SEC-SJT	PR _{gesamt}	PR _{weiblich}	PR _{männlich}
10	100	100	100
9	95	94	96
8	80	76	83
7	61	51	67
6	43	32	49
5	30	22	35
4	19	14	23
3	13	9	15
2	8	7	9
1	3	2	4
0	1	0	2

Anmerkung: SEC-SJT = Social-Emotional Competence Situational Judgement Test Rohwert; PR = Prozentrang.

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

2.4.12. Interpretationsbeispiele

Zur Veranschaulichung der Interpretation wird an die Beschreibung der Auswertung aus dem Abschnitt *Auswertung mittels Auswertungsbögen* (siehe S. 43) bzw. an die Beschreibung der Interpretation aus dem Abschnitt *Interpretation mit dem zugehörigen Profilbogen* (siehe S. 76) angeschlossen und die Interpretation für die Person mit einem unauffälligen Testergebnis fortgesetzt.

Für die Interpretation wurde zusätzlich zu dieser Person (siehe Abschnitt *Interpretationsbeispiel für ein unauffälliges Testergebnis*, S. 81) eine weitere Person mit einem auffälligen Testergebnis (siehe Abschnitt *Interpretationsbeispiel für ein auffälliges Testergebnis*, S. 84) aus der Gesamtstichprobe ausgewählt. Zur Wahrung ihrer Anonymität werden auch für diese Person keine Angaben zu Alter und Geschlecht gemacht. Für die Ermittlung des Normwerts wird daher für beide Interpretationsbeispiele auch die Gesamtstichprobe herangezogen.

2.4.12.1. Interpretationsbeispiel für ein unauffälliges Testergebnis

In Abbildung 16 (auf der nächsten Seite) ist der vollständig ausgefüllte Profilbogen einer Testperson mit einem unauffälligen Testergebnis abgebildet. Als Norm wurde die Gesamtstichprobe gewählt. Die *T*-Werte und Prozentränge wurden aus den Normtabellen entnommen (für den Fragebogen zur Selbsteinschätzung Tabelle A1 des Anhangs, siehe S. 136; für den Fragebogen zur Fremdeinschätzung Tabelle A4 des Anhangs, siehe S. 142; für den Situational Judgement Test Tabelle A7 des Anhangs, siehe S. 148) und auf Seite 3 des Profilbogens für die jeweilige Globalskala und die Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung in die Ergebnistabelle eingetragen.

Die Testperson erreicht mit einem PR von 80 auf der Globalskala Sozial-emotionale Kompetenz des Situational Judgement Test (SEC-SJT), einem *T*-Wert von 51 (PR von 56) auf der Globalskala Sozial-emotionale Kompetenz des Fragebogens zur Selbsteinschätzung (SEC-SR) sowie einem *T*-Wert von 55 (PR von 71) auf der Globalskala Sozial-emotionale Kompetenz des Fragebogens zur Fremdeinschätzung (SEC-OR) ein durchgehend in den Durchschnittsbereich einzuordnendes Testergebnis. Die globale sozial-emotionale Kompetenz ist damit bei multimethodaler Erfassung in objektiver Leistung, Selbstbild und Fremdbild unauffällig.

Abbildung 16

Profilbogen einer Testperson mit einem unauffälligen Testergebnis

PROFILBOGEN

Normierungsschritte

- Übertragen Sie die Rohwerte (Summen aus dem SEC-I-SR, SEC-I-OR und SEC-SJT Auswertungsbogen) der jeweiligen Skalen (Subskalen und Globalskalen) in die dafür vorgesehenen Zellen der **Ergebnistabelle**.
- Wählen Sie eine geeignete Norm (Gesamtstichprobe, Stichprobe weiblichen bzw. männlichen Geschlechts) und entnehmen Sie dann die zugehörigen T-Werte bzw. Prozentränge (PR) dem Testmanual.

Ergebnistabelle

Norm: Gesamt weiblich männlich

	Rohwert	T-Wert	PR		Rohwert	T-Wert	PR
Skala				Skala			
SEC-SJT	8		80				
SEC-SR	153	51	56	SEC-OR	147	55	71
EBE-SR	17	56	74	EBE-OR	14	52	60
ECO-SR	17	49	48	ECO-OR	17	50	52
PEA-SR	18	49	48	PEA-OR	17	49	48
PES-SR	21	50	52	PES-OR	19	51	56
REA-SR	16	48	44	REA-OR	14	50	52
RES-SR	16	57	77	RES-OR	16	66	95
SOB-SR	30	44	29	SOB-OR	34	57	77
SON-SR	18	50	52	SON-OR	16	52	60

3 von 5

PROFILBOGEN

Darstellungsschritte

- Übertragen Sie die T-Werte der Selbsteinschätzung (Skalen mit Endung -SR) aus der **Ergebnistabelle** in das **Profil für SEC-I**.
- Verbinden Sie die T-Werte auf den Subskalen (EBE, ECO, PEA, PES, REA, RES, SOB und SON) der Selbsteinschätzung mit einer Profillinie und kennzeichnen Sie diese.
- Wiederholen Sie diese Schritte für die Fremdeinschätzung (Skalen mit Endung -OR).
- Wählen Sie eine Vertrauenswahrscheinlichkeit, entnehmen Sie die Werte für die Spannweite des Vertrauensbereichs der jeweiligen Skala dem Testmanual und zeichnen Sie den Vertrauensbereich ein.

Profil für SEC-I

Vertrauenswahrscheinlichkeit: 95% (2-seitig)

4 von 5

Anmerkung. Testmaterialien einer Person mit einem unauffälligen Testergebnis.

Die T-Werte der Selbsteinschätzung (in blauer Farbe) und Fremdeinschätzung (in gelber Farbe) wurden leicht versetzt und in Farbe in das *Profil für SEC-I* übertragen, sodass ähnliche T-Werte besser unterschieden werden können. Die T-Werte auf den jeweiligen Subskalen wurden dann mit einer Profillinie verbunden und Vertrauensbereiche für eine 95% Vertrauenswahrscheinlichkeit bei zweiseitiger Fragestellung eingezeichnet (die Werte für die Spannweite der Vertrauensbereiche wurden Tabelle A9 des Anhangs entnommen, siehe S. 150).

Die Profillinie der Selbsteinschätzung verläuft weitestgehend geradlinig im Durchschnittsbereich. Die T-Werte auf allen Subskalen liegen im Durchschnittsbereich. Die 95%-Vertrauensbereiche der Subskalen der Selbsteinschätzung überschneiden sich durchgehend und überdecken jeweils gegenseitig die T-Werte auf den übrigen Subskalen. Es gibt keine Hinweise auf bedeutsame intraindividuelle Differenzen im Selbstbild.

Die Profillinie der Fremdeinschätzung verlässt auf der Subskala Regulation eigener Emotionen (RES-OR) den Durchschnittsbereich, die T-Werte auf den übrigen Subskalen liegen im Durchschnittsbereich. Der überdurchschnittliche T-Wert von 66 auf der Subskala RES-OR erweist sich bei Einbezug des 95%-Vertrauensbereichs als durchschnittlich bis

Interpretation

überdurchschnittlich, denn der 95%-Vertrauensbereich überschneidet den Durchschnittsbereich. Die 95%-Vertrauensbereiche der übrigen Subskalen überschneiden sich weitestgehend. Der 95%-Vertrauensbereich der Subskala RES-OR (T -Wert von 66, 95%-Vertrauensbereich [58, 74]) überschneidet sich nicht mit den 95%-Vertrauensbereichen der Subskala Kognitive Empathie (ECO-OR: T -Wert von 50, 95%-Vertrauensbereich [43, 57]), der Subskala Wahrnehmung der Emotionen anderer (PEA-OR: T -Wert von 49, 95%-Vertrauensbereich [41, 57]) und der Subskala Regulation der Emotionen anderer (REA-OR: T -Wert von 50, 95%-Vertrauensbereich [44, 56]). Um diese relative Stärke im Fremdbild weiter abzuklären, werden die kritischen Differenzen für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung herangezogen (entnommen aus Tabelle A10 des Anhangs, S. 151) und mit den empirischen Differenzen verglichen: Die kritische Differenz für eine statistisch bedeutsame intraindividuelle Differenz zwischen den Subskalen RES-OR und EBE-OR beträgt 11, die zwischen den Subskalen RES-OR und PES-OR 12, die zwischen den Subskalen RES-OR und SOB-OR 10 und die zwischen den Subskalen RES-OR und SON-OR 11. Bis auf die empirische Differenz zwischen den Subskalen RES-OR und SOB-OR ($66 - 57 = 9 \leq 10$), die nicht statistisch bedeutsam ausfällt, fallen die übrigen empirischen Differenzen statistisch bedeutsam aus (RES-OR und EBE-OR: $66 - 52 = 14 > 11$; RES-OR und PES-OR: $66 - 52 = 15 > 12$; RES-OR und SON-OR: $66 - 52 = 14 > 11$).

Die beiden Profillinien verlaufen weitestgehend parallel und überdecken sich teilweise. Die 95%-Vertrauensbereiche der Selbsteinschätzung und der Fremdeinschätzung auf inhaltsgleichen Subskalen überschneiden sich auf allen Subskalen. Auf den Subskalen RES-SR bzw. RES-OR und SOB-SR bzw. SOB-OR fallen Selbst- und Fremdeinschätzung am weitesten auseinander. Der 95%-Vertrauensbereich [47, 67] für den T -Wert von 57 auf der Subskala RES-SR der Selbsteinschätzung überdeckt den T -Wert von 66 auf der Subskala RES-OR der Fremdeinschätzung. Umgekehrt überdeckt der 95%-Vertrauensbereich [58, 74] für den T -Wert von 66 auf der Subskala RES-OR der Fremdeinschätzung nicht den T -Wert von 57 auf der Subskala RES-SR der Selbsteinschätzung. Für die Subskala SOB überdeckt weder der 95%-Vertrauensbereich [35, 53] der Selbsteinschätzung noch der 95%-Vertrauensbereich [50, 64] der Fremdeinschätzung den T -Wert von 57 der Fremdeinschätzung bzw. den T -Wert von 44 der Selbsteinschätzung. Um das augenfällige Auseinanderfallen der Selbst- und Fremdeinschätzung auf den Subskalen RES und SOB weiter abzuklären, werden die kritischen Differenzen für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung herangezogen (entnommen aus Tabelle A10 des Anhangs, siehe S. 151) und mit den empirischen Differenzen verglichen: Die kritische Differenz für eine statistisch bedeutsame intraindividuelle Differenz zwischen den Subskalen RES-SR und RES-OR beträgt 13, die zwischen den Subskalen SOB-SR und SOB-OR 12. Die empirische Differenz zwischen den Subskalen RES-SR und RES-OR

$(|57 - 66| = 9 \leq 13)$ fällt somit nicht statistisch bedeutsam aus, die zwischen den Subskalen SOB-SR und SOB-OR $(|44 - 57| = 13 > 12)$ fällt hingegen statistisch bedeutsam aus.

Anschließend an diese eingehenden Analysen kann das Testergebnis wie folgt interpretiert werden: Die Testperson weist insgesamt einen unauffälligen Umgang mit den Emotionen und Bedürfnissen anderer Personen wie auch der eigenen Person auf. Die objektive Leistung (Globalskala SEC-SJT), das globale Selbstbild (Globalskala SEC-SR) und das globale Fremdbild (Globalskala SEC-OR) fallen durchschnittlich aus. Dieses insgesamt unauffällige Testergebnis wird durch die Ergebnisse auf den Subskalen des Inventars bestätigt. Im Selbstbild weist die Testperson unauffällige homogene sozial-emotionale Kompetenzen auf, keine der Kompetenzen weicht von der durchschnittlichen Einschätzung ab. Im Fremdbild tritt die Fähigkeit regulierend auf das eigene emotionale Erleben einwirken und unangenehme Emotionen bewältigen zu können (Subskala RES-OR) in Relation zu den anderen Kompetenzen hervor, die Testperson wird hier als durchschnittlich bis überdurchschnittlich eingeschätzt. Selbstbild und Fremdbild stimmen weitestgehend überein, das Bemühen um den Aufbau und die Aufrechterhaltung positiver Beziehungen zu anderen Personen und die Übernahme von Verantwortung für das Wohlergehen anderer Personen (Subskala SOB) wird im Fremdbild höher eingeschätzt als von der Testperson selbst.

2.4.12.2. Interpretationsbeispiel für ein auffälliges Testergebnis

In Abbildung 17 (auf der nächsten Seite) ist der vollständig ausgefüllte Profilbogen einer Testperson mit einem auffälligen Testergebnis abgebildet. Als Norm wurde die Gesamtstichprobe gewählt. Die *T*-Werte und Prozentränge wurden aus den Normtabellen entnommen (für den Fragebogen zur Selbsteinschätzung Tabelle A1 des Anhangs, siehe S. 136; für den Fragebogen zur Fremdeinschätzung Tabelle A4 des Anhangs, siehe S. 142; für den Situational Judgement Test Tabelle A7 des Anhangs, siehe S. 148) und auf Seite 3 des Profilbogens für die jeweilige Globalskala und die Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung in die Ergebnistabelle eingetragen.

Die Testperson erreicht mit einem PR von 3 auf der Globalskala Sozial-emotionale Kompetenz des Situational Judgement Tests (SEC-SJT), einem *T*-Wert von 39 (PR von 15) auf der Globalskala Sozial-emotionale Kompetenz des Fragebogens zur Selbsteinschätzung (SEC-SR) sowie einem *T*-Wert von 24 (PR von 1) auf der Globalskala Sozial-emotionale Kompetenz des Fragebogens zur Fremdeinschätzung (SEC-OR) ein durchgehend in den unterdurchschnittlichen Bereich einzuordnendes Testergebnis. Der deskriptiv unterdurchschnittliche *T*-Wert von 39 auf der Globalskala SEC-SR erweist sich bei Einbezug des 95%-Vertrauensbereichs als unterdurchschnittlich bis durchschnittlich (der 95%-Vertrauensbereich überdeckt den Durchschnittsbereich). Die globale sozial-emotionale Kompetenz der Testperson ist damit bei multimethodaler Erfassung in objektiver Leistung, Selbstbild und Fremdbild auffällig.

Abbildung 17

Profilbogen einer Testperson mit einem auffälligen Testergebnis

PROFILBOGEN

Normierungsschritte

- Übertragen Sie die Rohwerte (Summen aus dem SEC-I-SR, SEC-I-OR und SEC-SJT Auswertungsbogen) der jeweiligen Skalen (Subskalen und Globalskalen) in die dafür vorgesehenen Zellen der **Ergebnistabelle**.
- Wählen Sie eine geeignete Norm (Gesamtstichprobe, Stichprobe weiblichen bzw. männlichen Geschlechts) und entnehmen Sie dann die zugehörigen T-Werte bzw. Prozentränge (PR) dem Testmanual.

Ergebnistabelle

Norm: Gesamt weiblich männlich

Skala	Rohwert	T-Wert	PR	Skala	Rohwert	T-Wert	PR
SEC-SJT	1		3				
SEC-SR	127	39	15	SEC-OR	74	24	1
EBE-SR	11	37	11	EBE-OR	5	26	1
ECO-SR	14	40	17	ECO-OR	9	26	1
PEA-SR	19	52	60	PEA-OR	10	27	1
PES-SR	14	36	9	PES-OR	13	33	5
REA-SR	19	57	77	REA-OR	5	28	2
RES-SR	16	57	77	RES-OR	7	32	4
SOB-SR	19	26	1	SOB-OR	16	29	2
SON-SR	15	41	20	SON-OR	9	33	5

3 von 5

PROFILBOGEN

Darstellungsschritte

- Übertragen Sie die T-Werte der Selbsteinschätzung (Skalen mit Endung -SR) aus der **Ergebnistabelle** in das **Profil für SEC-I**.
- Verbinden Sie die T-Werte auf den Subskalen (EBE, ECO, PEA, PES, REA, RES, SOB und SON) der Selbsteinschätzung mit einer Profillinie und kennzeichnen Sie diese.
- Wiederholen Sie diese Schritte für die Fremdeinschätzung (Skalen mit Endung -OR).
- Wählen Sie eine Vertrauenswahrscheinlichkeit, entnehmen Sie die Werte für die Spannweite des Vertrauensbereichs der jeweiligen Skala dem Testmanual und zeichnen Sie den Vertrauensbereich ein.

Profil für SEC-I

Vertrauenswahrscheinlichkeit: 95% (zweiseitig)

4 von 5

Anmerkung. Testmaterialien einer Person mit einem auffälligen Testergebnis.

Die *T*-Werte der Selbsteinschätzung (in blauer Farbe) und der Fremdeinschätzung (in gelber Farbe) wurden leicht versetzt und in Farbe in das *Profil für SEC-I* übertragen, sodass ähnliche *T*-Werte besser unterschieden werden können. Die *T*-Werte auf den jeweiligen Subskalen wurden dann mit einer Profillinie verbunden und Vertrauensbereiche für eine 95% Vertrauenswahrscheinlichkeit bei zweiseitiger Fragestellung eingezeichnet (die Werte für die Spannweite der Vertrauensbereiche wurden Tabelle A9 des Anhangs entnommen, siehe S. 150).

Die Profillinie der Selbsteinschätzung verläuft sprunghaft. Die *T*-Werte auf den Subskalen Kognitive Empathie (ECO-SR), Wahrnehmung der Emotionen anderer (PEA-SR), Regulation der Emotionen anderer (REA-SR), Regulation eigener Emotionen (RES-SR) und Beachtung sozialer Normen (SON-SR) liegen im Durchschnittsbereich. Die *T*-Werte auf den Subskalen Empathisches Verhalten (EBE-SR), Wahrnehmung eigener Emotionen (PES-SR) und Positive Beziehungsgestaltung (SOB-SR) liegen im unterdurchschnittlichen Bereich. Die 95%-Vertrauensbereiche der höchsten *T*-Werte auf den Subskalen PEA-SR (*T*-Wert von 52, 95%-Vertrauensbereich [42, 62]), REA-SR (*T*-Wert von 57, 95%-Vertrauensbe-

reich [49, 65]) und RES-SR (T -Wert von 57, 95%-Vertrauensbereich [47, 67]) überschneiden sich nicht mit dem 95%-Vertrauensbereich des niedrigsten T -Werts auf der Subskala SOB-SR (T -Wert von 26, 95%-Vertrauensbereich [17, 35]). Aufgrund dieser bereits deutlichen augenfälligen Hinweise auf bedeutsame intraindividuelle Differenzen im Selbstbild sind zusätzliche Vergleiche zwischen kritischen Differenzen und empirischen Differenzen unnötig.

Die Profillinie der Fremdeinschätzung verläuft deutlich geradliniger als die der Selbsteinschätzung und durchgehend im unterdurchschnittlichen Bereich. Die 95%-Vertrauensbereiche der Subskalen PES-OR, RES-OR und SON-OR überschneiden den Durchschnittsbereich und erweisen sich damit als unterdurchschnittlich bis durchschnittlich. Die 95%-Vertrauensbereiche sämtlicher Subskalen der Fremdeinschätzung überschneiden sich durchgehend und überdecken jeweils gegenseitig die T -Werte auf den übrigen Subskalen. Es gibt daher keine Hinweise auf bedeutsame intraindividuelle Differenzen zwischen jeweils zwei verschiedenen Subskalen im Fremdbild.

Die beiden Profillinien fallen weitestgehend auseinander. Im Selbstbild und Fremdbild zeigen sich ähnliche Einschätzungen auf den Subskalen PES und SOB, die 95%-Vertrauensbereiche überdecken gegenseitig die T -Werte auf den jeweils anderen Subskalen. Die Selbsteinschätzung und die Fremdeinschätzung fallen auf den Subskalen PEA, REA und RES sehr deutlich auseinander. Die 95%-Vertrauensbereiche dieser Subskalen überschneiden sich nicht. Auch auf den Subskalen EBE und ECO fallen Selbsteinschätzung und Fremdeinschätzung noch auseinander. Die 95%-Vertrauensbereiche überschneiden sich zwar, doch keiner der 95%-Vertrauensbereiche überdeckt einen der jeweils anderen T -Werte. Aufgrund dieser ebenfalls deutlich augenfälligen Hinweise auf bedeutsame interindividuelle Differenzen in Selbstbild und Fremdbild werden keine weiteren Vergleiche zwischen kritischen Differenzen und empirischen Differenzen in Betracht gezogen.

Anschließend an diese eingehenden Analysen kann das Testergebnis wie folgt interpretiert werden: Die Testperson weist insgesamt einen auffälligen Umgang mit den Emotionen und Bedürfnissen anderer Personen, wie auch der eigenen Person auf. Die objektive Leistung (Globalskala SEC-SJT), das globale Selbstbild (Globalskala SEC-SR) und das globale Fremdbild (Globalskala SEC-OR) fallen unterdurchschnittlich aus. Das insgesamt auffällige Testergebnis wird durch die Ergebnisse auf den Subskalen des Inventars weitestgehend bestätigt. Im Fremdbild weist die Testperson homogen niedrige sozial-emotionale Kompetenzen auf, keine der Kompetenzen weicht von dieser niedrigen Einschätzung ab. Im Selbstbild schätzt die Testperson ihre Kompetenzen heterogener ein. Insbesondere die Fähigkeiten regulierend auf das emotionale Erleben anderer Personen wie auch das eigene emotionale Erleben einwirken zu können (Subskalen REA-SR und RES-SR), treten in Relation zu den übrigen Kompetenzen hervor, werden jedoch nicht höher als durchschnittlich eingeschätzt. Im Selbstbild wird insbesondere – und in Übereinstimmung mit dem

Interpretation

Fremdbild – das Bemühen um den Aufbau und die Aufrechterhaltung positiver Beziehungen zu anderen Personen und die Übernahme von Verantwortung für das Wohlergehen anderer Personen (Subskala SOB) niedrig eingeschätzt.

3. Testtheoretische Fundierung

Dieses dritte Kapitel widmet sich der theoretischen Fundierung der diagnostischen Verfahren. Psychometrische Merkmale der Skalen werden untersucht, die Testgüte wird analysiert und bewertet. Die empirischen Prüfungen erfolgten anhand der Gesamtstichprobe von $N = 306$ Personen. Eine ausführliche Beschreibung der Rekrutierung sowie weiterer demografischer Merkmale der Gesamtstichprobe findet sich im Abschnitt *Beschreibung der Gesamtstichprobe* (siehe S. 120).

3.1. Psychometrische Merkmale

3.1.1. Itemanalyse

3.1.1.1. Analyse der Items des Fragebogens zur Selbsteinschätzung

Die Mittelwerte (M), Standardabweichungen (s), Schiefen (μ_3), Exzesse (μ_4) und part-whole korrigierten Trennschärfen ($r_{I(T-I)}$) der Items des Fragebogens zur Selbsteinschätzung sind für die Daten der Gesamtstichprobe ($N = 306$) in Tabelle A14 des Anhangs (siehe S. 155) angegeben. Abbildung A1 (siehe S. 162) bzw. Abbildung A2 (siehe S. 163) veranschaulichen die Häufigkeitsverteilungen der Skalenrohwerte der Globalskala SEC-SR bzw. der Subskalen.

Die Itemmittelwerte liegen im Bereich von $3.02 \leq M \leq 4.58$ mit einem Mittelwert von $M = 3.68$ deutlich über dem Mittelwert der Ratingskala. Die Schiefen liegen im Bereich von $-1.92 \leq \mu_3 \leq -0.03$ mit einem Mittelwert von $M = -0.57$. Damit weisen sämtlich Items der Selbsteinschätzung in der Gesamtstichprobe eine linksschiefe Verteilung und eine (relativ) hohe Itemschwierigkeit auf (anders als in der Umgangssprache weist ein Item vom Likert-Typ in der klassischen Testtheorie eine hohe Itemschwierigkeit auf, wenn es von einem Großteil der Testpersonen mit einem hohen Grad an Zustimmung beantwortet wurde). Die part-whole korrigierten Trennschärfen der Items für die jeweils zugehörige Subskala liegen im Bereich von $.24 \leq r_{I(T-I),sub} \leq .67$ mit einem Mittelwert von $M = .53$. Für die Globalskala SEC-SR liegen die part-whole korrigierten Trennschärfen der Items im Bereich von $.22 \leq r_{I(T-I),global} \leq .60$ mit einem Mittelwert von $M = .47$. Bei den meisten Items fällt die part-whole korrigierte Trennschärfe für die jeweils zugehörige Subskala höher aus als die part-whole korrigierte Trennschärfe für die Globalskala SEC-SR. Bei den Items der Subskalen ECO-SR und SOB-SR fallen die part-whole korrigierten Trennschärfen für die jeweils zugehörige Subskala und die Globalskala SEC-SR jedoch relativ ähnlich aus.

3.1.1.2. Analyse der Items des Fragebogens zur Fremdeinschätzung

Die Mittelwerte (M), Standardabweichungen (s), Schiefen (μ_3), Exzesse (μ_4) und part-whole korrigierten Trennschärfen ($r_{1(T-I)}$) der Items des Fragebogens zur Fremdeinschätzung sind für die Daten der Gesamtstichprobe ($N = 306$) in Tabelle A15 des Anhangs (siehe S. 157) angegeben. Abbildung A1 (siehe S. 162) bzw. Abbildung A3 (siehe S. 164) veranschaulichen die Häufigkeitsverteilungen der Skalenrohwerte der Globalskala SEC-OR bzw. der Subskalen.

Die Itemmittelwerte liegen im Bereich von $2.27 \leq M \leq 3.97$ mit einem Mittelwert von $M = 3.30$ näher am Mittelwert der Ratingskala, als die Itemmittelwerte des Fragebogens zur Selbsteinschätzung. Die Schiefen liegen im Bereich von $-0.70 \leq \mu_3 \leq 0.50$ mit einem Mittelwert von $M = -0.19$. Damit weist, anders als bei den Items des Fragebogens zur Selbsteinschätzung, nur ein Teil der Items des Fragebogens zur Fremdeinschätzung in der Gesamtstichprobe eine linksschiefe Verteilung auf. Die Itemschwierigkeiten fallen im Vergleich zum Fragebogen zur Selbsteinschätzung insgesamt niedriger aus. Die part-whole korrigierten Trennschärfen der Items für die jeweils zugehörige Subskala liegen im Bereich von $.38 \leq r_{1(T-I),sub} \leq .79$ mit einem Mittelwert von $M = .67$. Für die Globalskala SEC-OR liegen die part-whole korrigierten Trennschärfen der Items im Bereich von $.35 \leq r_{1(T-I),global} \leq .76$ mit einem Mittelwert von $M = .60$. Bei den meisten Items fällt die part-whole korrigierte Trennschärfe für die jeweils zugehörige Subskala höher aus als die part-whole korrigierte Trennschärfe für die Globalskala SEC-OR. Anders als beim Fragebogen zur Selbsteinschätzung, fallen beim Fragebogen zur Fremdeinschätzung für die Items der Subskala PEA-OR die part-whole korrigierten Trennschärfen für die zugehörige Subskala und die Globalskala SEC-OR relativ ähnlich aus.

3.1.1.3. Analyse der Items des Situational Judgement Tests

Die relativen Häufigkeiten der gewählten Antwortalternativen und part-whole korrigierten Trennschärfen der Items des SEC-SJT sind für die Daten der Gesamtstichprobe ($N = 306$) in Tabelle A16 des Anhangs (siehe S. 159) angegeben. Abbildung A1 (siehe S. 162) veranschaulicht die Häufigkeitsverteilung der Skalenrohwerte der Globalskala.

Die relativen Häufigkeiten der korrekten Antwort liegen im Bereich von $51.96\% \leq h \leq 76.14\%$ mit einem Mittelwert von $M = 63.43\%$. Damit weisen die Items des SEC-SJT eine mittlere bis hohe jedoch noch nicht als problematisch einzustufende Itemschwierigkeit auf. Umgangssprachlich würde man von „eher leichten“ Items sprechen (anders als in der Umgangssprache weist ein dichotomes Item in der klassischen Testtheorie eine hohe Itemschwierigkeit auf, wenn es von einem Großteil der Testpersonen korrekt gelöst wurde). Die part-whole korrigierten Trennschärfen der Items liegen für die Globalskala SEC-SJT im Bereich von $.07 \leq r_{1(T-I)} \leq .47$ mit einem Mittelwert von $M = .32$. Die

part-whole korrigierte Trennschärfe von Item 3 fällt mit $r_{1(T-1)} = .07$ deutlich niedriger aus als die der anderen Items. Da ein Kontexteffekt von Item 3 auf die Beantwortung der anderen Items nicht ausgeschlossen werden kann, Item 3 theoretisch relevante Aspekte sozial-emotionaler Kompetenzen zum Inhalt hat und ein Ausschluss von Item 3 zu keiner deutlichen Verbesserung der Testgüte beiträgt, wurde entschieden Item 3 beizubehalten und in die Auswertung miteinzubeziehen.

3.1.2. Skalenanalyse

3.1.2.1. Analyse der Globalskala und Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung

Mittelwerte (M), Standardabweichungen (s), Schiefen (μ_3) und Exzesse (μ_4) der Skalen des Fragebogens zur Selbsteinschätzung und der Skalen des Fragebogens zur Fremdeinschätzung werden in Tabelle A17 des Anhangs (siehe S. 160) gegenübergestellt. Die Mittelwerte der Subskalen des Fragebogens zur Selbsteinschätzung liegen im Bereich von $3.29 \leq M \leq 4.32$, die des Fragebogens zur Fremdeinschätzung im Bereich von $3.73 \leq M \leq 3.81$. Der Mittelwert der Globalskala SEC-SR beträgt $M = 3.69$, der Mittelwert der Globalskala SEC-OR $M = 3.30$. Die Schiefen der Subskalen liegen für den Fragebogen zur Selbsteinschätzung im Bereich von $-0.99 \leq \mu_3 \leq 0.09$ mit einem Mittelwert von $M = -0.46$, für den Fragebogen zur Fremdeinschätzung im Bereich von $-0.50 \leq \mu_3 \leq 0.03$ mit einem Mittelwert von $M = -0.18$. Die Schiefe der Globalskala SEC-SR beträgt $\mu_3 = -0.32$, die der Globalskala SEC-OR $\mu_3 = -0.17$. Damit weisen die Subskalen wie auch die Globalskala des Fragebogens zur Selbsteinschätzung und die Globalskala des Fragebogens zur Fremdeinschätzung eine linksschiefe Verteilung auf. Die Skalen des Fragebogens zur Selbsteinschätzung fallen insgesamt linksschiefer aus als die Skalen des Fragebogens zur Fremdeinschätzung. Die Skala SON fällt sowohl für den Fragebogen zur Selbsteinschätzung wie auch den Fragebogen zur Fremdeinschätzung am deutlichsten linksschief aus.

3.1.2.2. Analyse der Globalskala des Situational Judgement Tests

Mittelwert (M), Standardabweichung (s), Schiefe (μ_3) und Exzess (μ_4) der Globalskala des Situational Judgement Tests (SEC-SJT) sind in Tabelle A17 des Anhangs (siehe S. 160) angegeben. Der Mittelwert beträgt $M = 6.46$, die Schiefe $\mu_3 \leq -0.71$. Damit weist auch die Globalskala SEC-SJT eine linksschiefe Verteilung auf.

3.1.3. Skaleninterkorrelationen

Sämtliche Skaleninterkorrelationen finden sich in Tabelle A18 des Anhangs (siehe S. 161). Auf die Skaleninterkorrelationen wird im Rahmen der Multitrait-Multimethod-Analyse (siehe S. 101) ausführlich eingegangen.

3.2. Testgüte

3.2.1. Objektivität

Durchführung und Auswertung sollten, wenn möglich, von ein und derselben Person und nur in Ausnahmefällen von zwei verschiedenen Personen vorgenommen werden. Dieses Testmanual enthält detaillierte Hinweise zur Standardisierung der Erhebung (siehe Abschnitt *Durchführung*, S. 29) und Auswertung (siehe Abschnitt *Auswertung*, S. 34) des Fragebogens zur Selbsteinschätzung und des Situational Judgement Tests. Für die Erhebung gehören dazu unter anderem Hinweise zur Qualifikation der Testleitung, zu Anforderungen an die Testperson, zur Gestaltung der Erhebungssituation, zur Instruktion der Testperson und zum Umgang mit Fragen und Störungen. Für die Auswertung gehören dazu unter anderem Hinweise zur Qualifikation der auswertenden Person, zur Prüfung der Qualität der Bearbeitung und zur Auswertung mittels Auswertungsbögen. Die Auswertung mittels Auswertungsbögen wird ausführlich anhand von Abbildungen erläutert. Darüber hinaus wird eine Auswertungsdatei (Microsoft® Excel® Arbeitsmappe) zur computerbasierten Auswertung zur Verfügung gestellt.

Das Testmanual enthält auch detaillierte Hinweise zur Interpretation (siehe Abschnitt *Interpretation*, S. 50). Dazu gehören Hinweise zur Qualifikation der interpretierenden Person, zur Transformation von Rohwerten in standardisierte Werte, zur Interpretation einzelner Skalen sowie einer Profillinie und zu Ergebnisrückmeldungen. Die Interpretation mittels Profilbogen wird ausführlich anhand von Abbildungen erläutert und mit zwei Interpretationsbeispielen veranschaulicht. Zur Förderung der Interpretationsobjektivität werden Rohwerte in *T*-Werte transformiert, für die auch die entsprechenden Prozentränge angegeben sind. *T*-Werte und Prozentränge werden bei Nutzung der Auswertungsdatei automatisch ausgegeben. Für die Interpretation werden verschiedene Vertrauensbereiche, kritische Differenzen für verschiedene Fragestellungen sowie ausführliche inhaltliche Beschreibungen der Skalen zur Verfügung gestellt.

3.2.1.1. Bewertung der Objektivität

Ist bzw. sind die mit Durchführung und Auswertung betraute(n) Person(en) mit den hierfür relevanten Inhalten des Testmanuals vertraut, wurde der Umgang mit den Testmaterialien eingeübt und werden sämtliche Hinweise beachtet, ist davon auszugehen, dass Durchführungs- und Auswertungsobjektivität gewährleistet sind. Wird computerbasiert ausgewertet, ist die Auswertungsobjektivität nur noch durch Flüchtigkeitsfehler bei der Übertragung der numerischen Werte der Antworten auf die einzelnen Items in die Auswertungsdatei gefährdet. Die Objektivität der Durchführung und der Auswertung ist damit gewährleistet. Erfolgt die Interpretation durch Fachpersonal mit fundierter psychologischer Ausbildung unter Berücksichtigung der relevanten Inhalte des Testmanuals sollte auch die Objektivität der Interpretation gewährleistet sein.

3.2.2. Reliabilität

Die Reliabilität bezeichnet die Messgenauigkeit, also den Grad der Genauigkeit, mit dem ein erfasstes Merkmal gemessen wird. Es ist gängiger Standard, die Reliabilität mit verschiedenen Methoden zu prüfen, unter anderem der Konsistenzanalyse und der Testwiederholungsmethode. Je nach verwendeter Methode werden (geringfügig) voneinander abweichende Aspekte der Messgenauigkeit betrachtet. Für das Inventar und den Situational Judgement Test wurde die Reliabilität anhand der Konsistenzanalyse geprüft (siehe Abschnitt *Konsistenzanalyse*, S. 94). Die Konsistenzanalyse prüft die Messgenauigkeit einer einmaligen Messung. Die interne Konsistenz trifft damit das, was im eigentlichen Sinn unter Messgenauigkeit zu verstehen ist, am ehesten. Die Testwiederholungsmethode prüft die Stabilität eines gemessenen Merkmals in Abhängigkeit von der zwischen zwei Messungen vergangenen Zeit und setzt eine angemessene Messgenauigkeit der beiden Messungen voraus. Da die Testwiederholungsmethode nicht angewendet wurde, ist die Stabilität der mit Inventar und Situational Judgement Test erfassten Merkmale bislang unklar (siehe Abschnitt *Testwiederholungsmethode*, S. 96). Auf Basis der internen Konsistenz wurde auch die Profilreliabilität nach Lienert (1956) berechnet (siehe Abschnitt *Profilreliabilität*, S. 96).

3.2.2.1. Konsistenzanalyse

Die interne Konsistenz sämtlicher Skalen bzw. erfasster Merkmale wurde anhand von Cronbachs Alpha (Cronbach, 1951; Cronbach & Shavelson, 2004) für die gesamte Stichprobe sowie die geschlechtsspezifischen Substichproben ermittelt und ist in Tabelle 1 (auf der nächsten Seite) angegeben. Für die Prüfung auf Homogenität der Cronbachs Alpha Koeffizienten zwischen den beiden geschlechtsspezifischen Subgruppen wurde die *F*-Statistik

als Quotient der Unreliabilitäten berechnet. Für den zugehörigen p -Wert wurden die Freiheitsgrade der F -Statistik nach Feldt (1969) und Feldt und Kim (2006) bestimmt.

Tabelle 1

Cronbachs Alpha Werte für die Skalen des Situational Judgement Tests (SEC-SJT), des Fragebogens zur Selbsteinschätzung (SEC-I-SR) und des Fragebogens zur Fremdeinschätzung (SEC-I-OR) für die Gesamtstichprobe und die geschlechtsspezifischen Substichproben sowie Statistiken der Prüfung auf Homogenität zwischen den Substichproben

Skala	Itemanzahl	α			F	p
		gesamt ($N = 306$)	weiblich ($n = 111$)	männlich ($n = 195$)		
SEC-SJT	10	0.65	0.62	0.65	1.11	.285
SEC-SR	41	0.92	0.93	0.92	1.01	.475
SEC-OR	41	0.96	0.96	0.96	1.14	.230
EBE-SR	4	0.80	0.78	0.80	1.09	.325
ECO-SR	5	0.75	0.77	0.74	1.12	.285
PEA-SR	5	0.74	0.68	0.77	1.37	.047
PES-SR	5	0.77	0.81	0.74	1.32	.075
REA-SR	5	0.82	0.83	0.81	1.15	.236
RES-SR	4	0.75	0.76	0.72	1.18	.204
SOB-SR	9	0.77	0.74	0.78	1.17	.187
SON-SR	4	0.68	0.68	0.68	1.03	.439
EBE-OR	4	0.87	0.88	0.85	1.24	.144
ECO-OR	5	0.87	0.87	0.87	1.01	.488
PEA-OR	5	0.83	0.84	0.83	1.09	.320
PES-OR	5	0.81	0.81	0.82	1.09	.329
REA-OR	5	0.91	0.93	0.90	1.43	.032
RES-OR	4	0.84	0.84	0.83	1.04	.430
SOB-OR	9	0.88	0.89	0.87	1.16	.211
SON-OR	4	0.86	0.90	0.83	1.71	.004

Anmerkung. α = Cronbachs Alpha; F = F -Statistik; p = p -Wert der F -Statistik für Freiheitsgrade nach Feldt (1969) und Feldt und Kim (2006); SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen; -SJT bzw. -SR bzw. -OR = Leistung bzw. Selbstbild bzw. Fremdbild.

Tabelle 1 kann entnommen werden, dass sozial-emotionale Kompetenz in den Stichproben auf der Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und auf der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) sehr reliabel erfasst werden kann, auf der Globalskala des Situational Judgement Tests (SEC-SJT) hingegen nur eingeschränkt reliabel. Die Reliabilität der Subskalen des Fragebogens zur Selbsteinschätzung fällt gemischt aus, die Subskalen weisen in der Gesamtstichprobe und den geschlechtsspezifischen Substichproben weitestgehend angemessene Reliabilitäten auf, die von $\alpha = .82$ (Subskala REA-SR) bis $\alpha = .68$ (Subskala SON-SR) reichen und im Mittel $M = .76$ betragen. Die Reliabilität der Subskalen des Fragebogens zur Fremdeinschätzung fällt im Vergleich zum Fragebogen zur Selbsteinschätzung einheitlicher aus, die Subskalen weisen in der Gesamtstichprobe und den geschlechtsspezifischen Substichproben gute Reliabilitäten auf, die von $\alpha = .91$ (Subskala REA-OR) bis $\alpha = .81$ (Subskala PES-OR) reichen und im Mittel $M = .86$ betragen. Die drei Subskalen PEA-SR ($\alpha_{weiblich} = .68$ vs. $\alpha_{männlich} = .77$, $F(97,135) = 1.37$, $p = .047$), REA-OR ($\alpha_{weiblich} = .93$ vs. $\alpha_{männlich} = .90$, $F(135,97) = 1.43$, $p = .032$) und SON-OR ($\alpha_{weiblich} = .90$ vs. $\alpha_{männlich} = .83$, $F(123,93) = 1.71$, $p = .004$) weisen in den geschlechtsspezifischen Subgruppen verschiedene Reliabilitäten auf.

3.2.2.2. Testwiederholungsmethode

Für die vorgesehene Prüfung der Stabilität der erfassten Merkmale konnte bis zum Zeitpunkt der Veröffentlichung dieses Testmanuals keine Stichprobe erhoben werden. Die Prüfung der Stabilität anhand einer Stichprobe von mindestens 100 Personen über ein Zeitintervall von mindestens drei Monaten steht noch aus.

3.2.2.3. Profilreliabilität

Die Profilreliabilität nach Lienert (1956) ist ein Maß der Messgenauigkeit eines Testprofils und immer dann von Bedeutung, wenn Profillinien interpretiert werden (siehe Abschnitt *Interpretation einer Profillinie*, S. 61). Die Interpretation einer Profillinie ist nur dann sinnvoll, wenn Differenzen zwischen einzelnen Werten eines Profils bedeutend und nicht zufalls- bzw. messfehlerbedingt sind (siehe Abschnitt *Interpretation von Differenzen*, S. 63). Die Profilreliabilität hängt sowohl von den Reliabilitäten der einfließenden Skalen wie auch deren Interkorrelationen ab und unterscheidet sich immer dann vom Mittelwert der Reliabilitäten der Skalen, wenn zwischen den Skalen Korrelationen bestehen (Lienert, 1956; Lienert & Ratz, 1998).

Die Profilreliabilität wurde für die Profillinie der Subskalen der Selbst- und Fremdeinschätzung auf Basis der internen Konsistenz berechnet. Die Profillinie der Subskalen der Selbsteinschätzung weist in der Gesamtstichprobe eine Profilreliabilität von $\rho_{\text{profilPt}} = .58$ auf, die Profillinie der Subskalen der Fremdeinschätzung eine Profilreliabilität von

$\rho_{tt} = .69$. Damit fällt die Reliabilität der Profillinie des Fragebogens zur Fremdeinschätzung höher aus als die des Fragebogens zur Selbsteinschätzung, sodass für diese gilt, dass bereits geringere Differenzen von Bedeutung sein können.

3.2.2.4. Bewertung der Reliabilität

Die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) weisen eine sehr gute Messgenauigkeit auf. Auch die Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung weisen eine weitestgehend gute Messgenauigkeit auf. Es liegen keine gravierenden Hinweise auf Unterschiede in der Messgenauigkeit für weibliche und männliche Testpersonen vor. Die Profilreliabilitäten für den Fragebogen zur Selbsteinschätzung und den Fragebogen zur Fremdeinschätzung fallen in einen angemessenen Bereich. Die Globalskala des Situational Judgement Tests weist hingegen nur eingeschränkte Messgenauigkeit auf. Die Prüfung der Stabilität aller mit den diagnostischen Verfahren erfassten Merkmale steht noch aus und sollte nachgeholt werden. Es wird empfohlen, in erster Linie die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und des Fragebogens zur Fremdeinschätzung (SEC-OR) zu interpretieren (siehe Abschnitt *Allgemeine Hinweise zur Interpretation*, S. 51). Bei der Interpretation der Globalskala des Situational Judgement Tests (SEC-SJT) sollte die eingeschränkte Reliabilität berücksichtigt werden.

3.2.3. Validität

3.2.3.1. Faktorielle Validität

Für die Analyse der faktoriellen Struktur der Globalskala des Situational Judgement Tests sowie der Globalskala und Subskalen des Fragebogens zur Selbsteinschätzung bzw. des Fragebogens zur Fremdeinschätzung wurden Hauptkomponentenanalysen für die Daten der Gesamtstichprobe ($N = 306$) berechnet. Für die Globalskala des Situational Judgement Tests (SEC-SJT) sowie die Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung (EBE, ECO, PEA, PES, REA, RES, SOB und SON) wurden die Hauptkomponentenanalysen auf Basis der jeweils theoriegeleitet entwickelten skalenzugehörigen Items berechnet. Für die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) wurden die theoriegeleitet entwickelten skalenzugehörigen Items zu jeweils einem Itembündel zusammengefasst und die Hauptkomponentenanalysen auf Basis der Itembündel berechnet. Bei den Hauptkomponentenanalysen wurde so vorgegangen, dass zunächst alle Komponenten mit Eigenwerten > 1.00 extrahiert wurden. Konnte für die

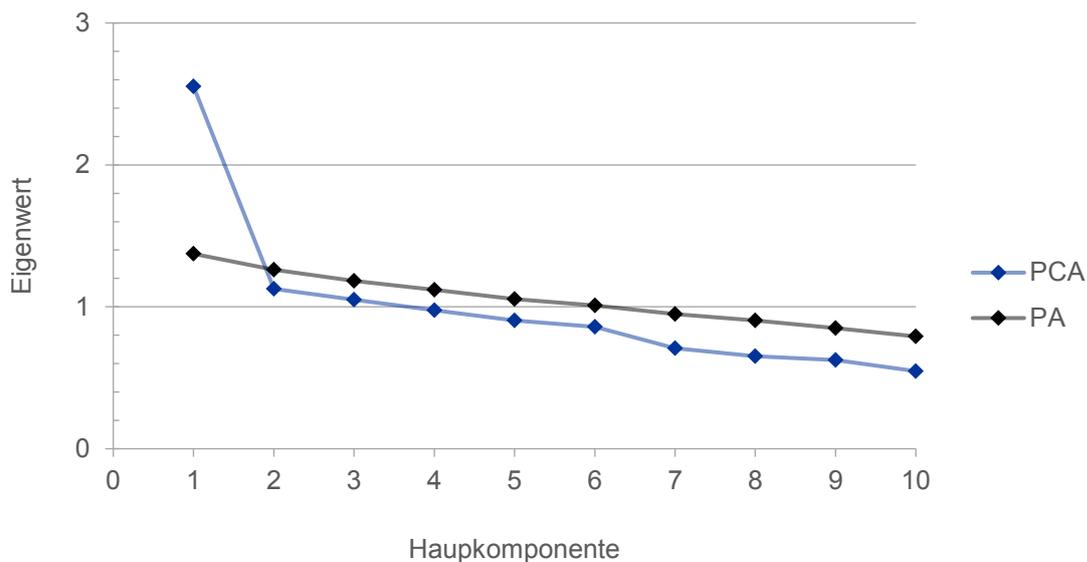
jeweilige Skala auf diese Weise mehr als eine Komponente extrahiert werden, wurde der jeweilige Eigenwertverlauf der Skala in einer Parallelanalyse nach Horn (1965) dem Eigenwertverlauf von 500 zufälligen Datensätzen (Obergrenze des 95%-Konfidenzintervalls) mit einer Stichprobengröße von $N = 306$ und der jeweils gleichen Anzahl an Items bzw. Itembündeln gegenübergestellt. Für die Parallelanalyse wurde die SPSS Syntax von O'Connor (2000) verwendet. Wurde durch die Parallelanalyse die Extraktion von mehr als einer Komponente nahegelegt, ist die entsprechende Anzahl an Komponenten extrahiert worden. Für jede Hauptkomponentenanalyse wird der Anteil erklärter Varianz an der Gesamtvarianz der entsprechenden Skala berichtet.

Struktur des Situational Judgement Tests (SEC-SJT)

Die Hauptkomponentenanalyse der Globalskala SEC-SJT mit 10 Items ergab drei Eigenwerte > 1.00 (2.55, 1.13 und 1.05). Durch den Eigenwertverlauf von 500 zufälligen Datensätzen aus der Parallelanalyse (1.37, 1.26, 1.18, ...) wurde die Extraktion einer Komponente nahegelegt. Die Lösung mit einer Hauptkomponente erklärt 25.5% der Gesamtvarianz der Globalskala SEC-SJT. Der Eigenwertverlauf der Hauptkomponentenanalyse für die Globalskala SEC-SJT bzw. der Parallelanalyse ist in Abbildung 18 dargestellt.

Abbildung 18

Parallelanalyse der (Subskalen-)Itembündel für die Globalskala des Situational Judgement Tests (SEC-SJT)



Anmerkung. $N = 306$. PCA = Eigenwertverlauf der Hauptkomponentenanalyse; PA = Eigenwertverlauf der Parallelanalyse.

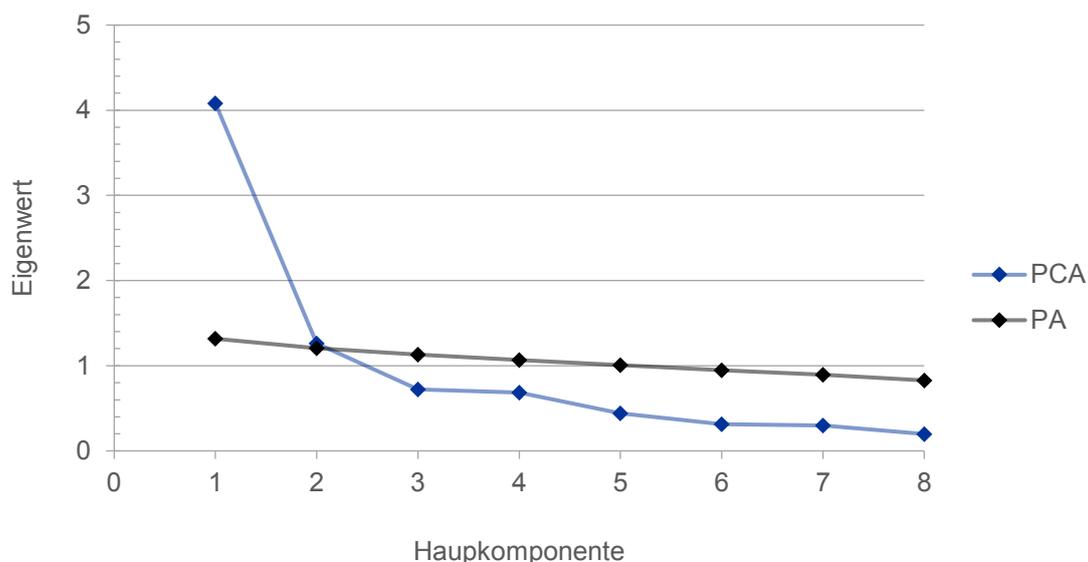
Struktur des Fragebogens zur Selbsteinschätzung (SEC-I-SR)

Die separaten Hauptkomponentenanalysen für die Subskalen EBE-SR mit 4 Items, ECO-SR mit 5 Items, PEA-SR mit 5 Items, PES-SR mit 5 Items, REA-SR mit 5 Items, RES-SR mit 4 Items und SON-SR mit 4 Items des Fragebogens zur Selbsteinschätzung ergaben jeweils nur einen einzigen Eigenwert > 1.00 . Die Lösung mit einer Hauptkomponente erklärt jeweils 62.5% der Gesamtvarianz der Subskala EBE-SR, 50.5% der Gesamtvarianz der Subskala ECO-SR, 49.3% der Gesamtvarianz der Subskala PEA-SR, 53.0% der Gesamtvarianz der Subskala PES-SR, 57.9% der Gesamtvarianz der Subskala REA-SR, 57.2% der Gesamtvarianz der Subskala RES-SR und 52.3% der Gesamtvarianz der Subskala SON-SR.

Die Hauptkomponentenanalyse für die Subskala SOB-SR mit 9 Items ergab zwei Eigenwerte > 1.00 (3.39 und 1.12). Durch den Eigenwertverlauf von 500 zufälligen Datensätzen aus der Parallelanalyse (1.34, 1.23, ...) wurde jedoch die Extraktion nur einer Komponente nahegelegt. Die Lösung mit einer Hauptkomponente erklärt 37.6% der Gesamtvarianz der Subskala SOB-SR.

Die Hauptkomponentenanalyse für die Globalskala SEC-SR mit 8 Itembündeln ergab zwei Eigenwerte > 1.00 (4.08 und 1.26). Durch den Eigenwertverlauf von 500 zufälligen Datensätzen aus der Parallelanalyse (1.31, 1.21, ...) wurde die Extraktion von zwei Komponenten nahegelegt.

Parallelanalyse der (Subskalen-)Itembündel für die Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR)



Anmerkung. $N = 306$. PCA = Eigenwertverlauf der Hauptkomponentenanalyse; PA = Eigenwertverlauf der Parallelanalyse.

Die Lösung mit zwei Hauptkomponenten erklärt 66.8% der Gesamtvarianz. Der Eigenwertverlauf der Hauptkomponentenanalyse für die Globalskala SEC-SR bzw. der Parallelanalyse ist in 0 dargestellt.

Die beiden extrahierten Komponenten wurden anschließend einer schiefwinkligen Oblimin-Rotation unterzogen. Werden Ladungen $< .30$ als vernachlässigbar angesehen, ergaben sich eindeutige Ladungen auf Komponente 1 für die Itembündel EBE-SR, SOB-SR, REA-SR, ECO-SR, PEA-SR und SON-SR (mit Ladungen in Höhe von .92, .88, .87, .69, .69 und .58) und auf Komponente 2 für die Itembündel RES-SR und PES-SR (mit Ladungen in Höhe von .83 und .79). Die beiden Komponenten korrelierten zu $r = .30$.

Um den Fragebogen zur Selbsteinschätzung möglichst parallel zum Fragebogen zur Fremdeinschätzung zu konstruieren, wurde auch die Extraktion nur einer Hauptkomponente analysiert. Eine Lösung mit nur einer Hauptkomponente erklärt noch 51.0% der Gesamtvarianz. Aufgrund des relativ geringen Eigenwerts der zweiten Hauptkomponente (der Eigenwert liegt nur knapp über dem Eigenwert der Parallelanalyse), der Korrelation der beiden Hauptkomponenten und einer Varianzaufklärung von knapp über 50.0%, wurde zugunsten der Lösung mit nur einer Hauptkomponente entschieden und die inhaltsähnliche Interpretation der Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-SR) höher gewichtet als der zusätzliche Informationsgewinn.

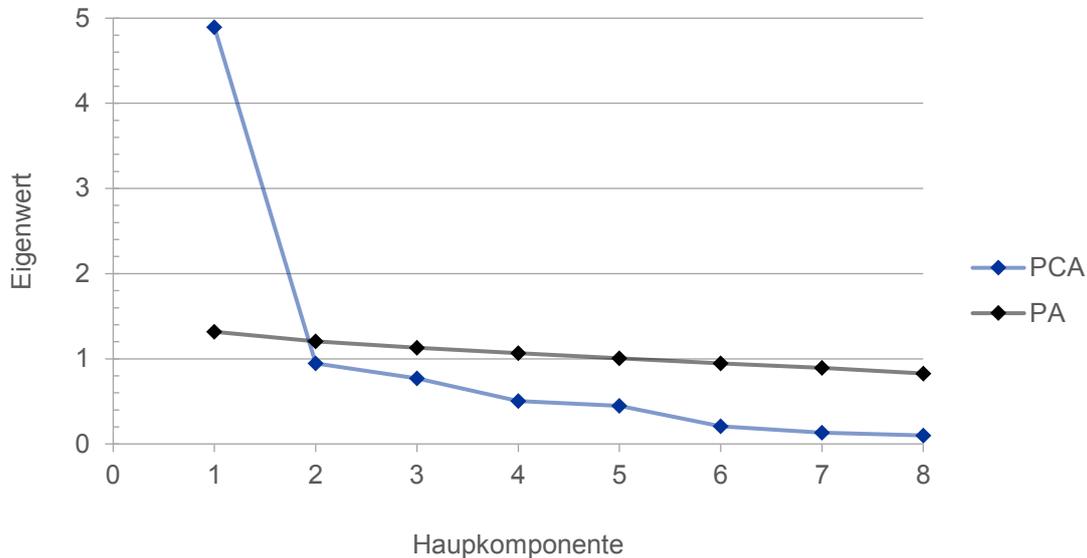
Struktur des Fragebogens zur Fremdeinschätzung (SEC-I-OR)

Die separaten Hauptkomponentenanalysen für die Subskalen EBE-OR mit 4 Items, ECO-OR mit 5 Items, PEA-OR mit 5 Items, PES-OR mit 5 Items, REA-OR mit 5 Items, RES-OR mit 4 Items, SON-OR mit 4 Items und SOB-OR mit 9 Items des Fragebogens zur Fremdeinschätzung ergaben jeweils nur einen Eigenwert > 1.00 . Die Lösung mit einer Hauptkomponente erklärt jeweils 71.3% der Gesamtvarianz der Subskala EBE-OR, 66.6% der Gesamtvarianz der Subskala ECO-OR, 60.2% der Gesamtvarianz der Subskala PEA-OR, 57.7% der Gesamtvarianz der Subskala PES-OR, 74.0% der Gesamtvarianz der Subskala REA-OR, 68.1% der Gesamtvarianz der Subskala RES-OR, 51.5% der Gesamtvarianz der Subskala SOB-OR und 70.3% der Gesamtvarianz der Subskala SON-OR.

Die Hauptkomponentenanalyse für die Globalskala SEC-OR mit 8 Itembündeln ergab nur einen Eigenwert > 1.00 (4.89). Die Lösung mit einer Hauptkomponente erklärt 61.2% der Gesamtvarianz der Globalskala SEC-OR. Der Eigenwertverlauf der Hauptkomponentenanalyse für die Globalskala SEC-OR bzw. der Parallelanalyse ist in Abbildung 19 (auf der nächsten Seite) dargestellt.

Abbildung 19

Parallelanalyse der (Subskalen-)Itembündel für die Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR)



Anmerkung. $N = 306$. PCA = Eigenwertverlauf der Hauptkomponentenanalyse; PA = Eigenwertverlauf der Parallelanalyse.

Bewertung der faktoriellen Validität

Die (explorativen) Analysen der faktoriellen Struktur der Skalen des Situational Judgement Tests, des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung deuten darauf hin, dass (weitestgehend) homogene Merkmale erfasst werden und die eindimensionale Struktur der Globalskalen und Subskalen (weitestgehend) angemessen erscheint. Durch die eindimensionale Struktur können die parallel konstruierten Verfahren inhaltlich ähnlich interpretiert werden. Die Struktur der Skalen der drei diagnostischen Verfahren sollte noch gemeinsam in einer größeren Stichprobe mittels konfirmatorischer Faktorenanalyse überprüft werden.

3.2.3.2. Multitrait-Multimethod-Analyse

Konvergente und divergente Validität der Skalen wurde anhand einer Multitrait-Multimethod-Analyse geprüft. Bei einer Multitrait-Multimethod-Analyse wird angenommen, dass psychologische Tests Merkmale nicht unabhängig von der Messmethode erfassen und damit erfasste Merkmale eine Verbindung aus Merkmal und Messmethode darstellen (Campbell & Fiske, 1959).

Nach Campbell und Fiske (1959) gilt, dass konvergente Validität vorliegt, wenn (a) gleiche Merkmale, die mit verschiedenen Messmethoden erfasst wurden, korrelieren. Divergente Validität hingegen liegt vor, wenn die Korrelation gleicher Merkmale, die mit verschiedenen Messmethoden erfasst wurden, höher ausfällt als die Korrelation von (b) verschiedenen Merkmalen, die mit der gleichen Messmethode erfasst wurden, und (c) verschiedenen Merkmalen, die mit verschiedenen Messmethoden erfasst wurden. Die auf das eigentliche Merkmal zurückführbare Ähnlichkeit sollte also größer sein als die auf die Messmethode zurückführbare Ähnlichkeit. Die Korrelation verschiedener Merkmale sollte (d) unabhängig davon, ob die verschiedenen Merkmale mit der gleichen Messmethode (Multitrait-Monomethod-Ansatz) oder mit verschiedenen Messmethoden (Multitrait-Multimethod-Ansatz) erfasst wurden, in ihrem Muster ähnlich sein. Die für die Multitrait-Multimethod-Analyse analysierten Korrelationen finden sich in Tabelle A18 des Anhangs (siehe S. 161).

Korrelationen zwischen den Globalskalen SEC-SJT, SEC-SR und SEC-OR

Die drei Globalskalen SEC-SR, SEC-OR und SEC-SJT erfassen das gleiche Merkmal, globale sozial-emotionale Kompetenz, mit verschiedenen Messmethoden (Fragebogen zur Selbsteinschätzung, Fragebogen zur Fremdeinschätzung und Situational Judgement Test). Sämtliche mögliche Korrelationen zwischen den drei Globalskalen (3 von 3) sind statistisch signifikant und liegen im Bereich von $.20 \leq r \leq .27$ mit einem Mittelwert von $M = .23$ und p -Werten $\leq .001$.

Korrelationen zwischen der Globalskala SEC-SJT und den SEC-I-SR Subskalen

Von sämtlichen möglichen Korrelationen zwischen der Globalskala des Situational Judgement Tests (SEC-SJT) und den acht Subskalen des Fragebogens zur Selbsteinschätzung sind 87.5% (7 von 8) statistisch signifikant. Ohne die nicht statistisch signifikante Korrelation zwischen der Globalskala SEC-SJT mit der Subskala RES-SR ($r(304) = -.07, p = .235$) liegen die Korrelationen der Globalskala SEC-SJT mit den Subskalen des Fragebogens zur Selbsteinschätzung im Bereich von $.14 \leq r(304) \leq .27$ mit einem Mittelwert von $M = .19$ und p -Werten $\leq .014$ bzw. mit der nicht statistisch signifikanten Korrelation im Bereich von $-.07 \leq r(304) \leq .27$ mit einem Mittelwert von $M = .16$ und p -Werten $\leq .235$.

Korrelationen zwischen der Globalskala SEC-SJT und den SEC-I-OR Subskalen

Von sämtlichen möglichen Korrelation zwischen der Globalskala des Situational Judgement Tests (SEC-SJT) und den acht Subskalen des Fragebogens zur Fremdeinschätzung sind alle (8 von 8) statistisch signifikant. Die Korrelationen der Globalskala SEC-SJT mit den Subskalen des Fragebogens zur Fremdeinschätzung liegen im Bereich von $.26 \leq r(304) \leq .42$ mit einem Mittelwert von $M = .34$ und p -Werten $\leq .001$.

Korrelationen zwischen der Globalskala SEC-SR und den SEC-I-SR Subskalen

Von sämtlichen möglichen Korrelation zwischen der Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und den acht Subskalen des Fragebogens zur Selbsteinschätzung sind alle (8 von 8) statistisch signifikant. Die Korrelationen der Globalskala SEC-SR mit den Subskalen des Fragebogens zur Selbsteinschätzung liegen im Bereich von $.43 \leq r(304) \leq .84$ mit einem Mittelwert von $M = .70$ und p -Werten $\leq .001$.

Korrelationen zwischen der Globalskala SEC-SR und den SEC-I-OR Subskalen

Von sämtlichen möglichen Korrelation zwischen der Globalskala des Fragebogens zur Selbsteinschätzung (SEC-SR) und den acht Subskalen des Fragebogens zur Fremdeinschätzung sind 87.5% (7 von 8) statistisch signifikant. Ohne die nicht statistisch signifikante Korrelation zwischen der Globalskala SEC-SR mit der Subskala RES-OR ($r(304) = .09, p = .136$) liegen die Korrelationen der Globalskala SEC-SR mit den Subskalen des Fragebogens zur Fremdeinschätzung im Bereich von $.14 \leq r(304) \leq .22$ mit einem Mittelwert von $M = .17$ und p -Werten $\leq .018$ bzw. mit der nicht statistisch signifikanten Korrelation im Bereich von $.09 \leq r(304) \leq .22$ mit einem Mittelwert von $M = .16$ und p -Werten $\leq .136$.

Korrelationen zwischen der Globalskala SEC-OR und den SEC-I-SR Subskalen

Von sämtlichen möglichen Korrelation zwischen der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) und den acht Subskalen des Fragebogens zur Selbsteinschätzung sind 75.0% (6 von 8) statistisch signifikant. Ohne die statistisch nicht signifikanten Korrelationen zwischen der Globalskala SEC-OR mit der Subskala REA-SR ($r(304) = .07, p = .222$) und zwischen der Globalskala SEC-OR mit der Subskala RES-SR ($r(304) = .08, p = .168$) liegen die Korrelationen der Globalskala SEC-OR mit den Subskalen des Fragebogens zur Selbsteinschätzung im Bereich von $.12 \leq r(304) \leq .22$ mit einem Mittelwert von $M = .17$ und p -Werten $\leq .044$ bzw. mit den nicht statistisch signifikanten Korrelation im Bereich von $.07 \leq r(304) \leq .22$ mit $M = .15$ und p -Werten $\leq .222$.

Korrelationen zwischen der Globalskala SEC-OR und den SEC-I-OR Subskalen

Von sämtlichen möglichen Korrelation zwischen der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-OR) und den acht Subskalen des Fragebogens zur Fremdeinschätzung sind alle (8 von 8) statistisch signifikant. Die Korrelationen der Globalskala SEC-OR mit den Subskalen des Fragebogens zur Fremdeinschätzung liegen im Bereich von $.57 \leq r(304) \leq .89$ mit einem Mittelwert von $M = .78$ und p -Werten $\leq .001$.

Korrelationen zwischen inhaltsgleichen SEC-I-SR und SEC-I-OR Subskalen

Der Fragbogen zur Selbsteinschätzung und der Fragebogen zur Fremdeinschätzung erfassen auf den Subskalen gleiche Merkmale mit verschiedenen Messmethoden (beispielsweise Empathisches Verhalten auf der Subskala EBE-SR im Selbstbild und der Subskala EBE-OR im Fremdbild). Von sämtlichen möglichen Korrelationen zwischen inhaltsgleichen, jedoch mit verschiedenen Messmethoden erfassten Subskalen sind 87.5% (7 von 8) statistisch signifikant. Ohne die nicht statistisch signifikante Korrelation zwischen der Subskalen REA-SR mit der Subskala REA-OR ($r(304) = .10, p = .090$) liegen die Korrelationen inhaltlich übereinstimmender Subskalen im Bereich von $.13 \leq r(304) \leq .29$ mit einem Mittelwert von $M = .20$ und p -Werten $\leq .026$ bzw. mit der nicht statistisch signifikanten Korrelation im Bereich von $.10 \leq r(304) \leq .29$ mit einem Mittelwert von $M = .19$ und p -Werten $\leq .090$.

Korrelationen zwischen nicht inhaltsgleichen SEC-I-SR und SEC-I-OR Subskalen

Der Fragbogen zur Selbsteinschätzung und der Fragebogen zur Fremdeinschätzung erfassen auf den Subskalen auch verschiedene Merkmale mit verschiedenen Messmethoden (beispielsweise Empathisches Verhalten auf der Subskala EBE-SR im Selbstbild und die Beachtung sozialer Normen auf der Subskala SON-OR im Fremdbild). Von sämtlichen möglichen Korrelationen zwischen nicht inhaltsgleichen, mit verschiedenen Messmethoden erfassten Subskalen sind 48.2% (27 von 56) statistisch signifikant. Die statistisch signifikanten Korrelationen zwischen inhaltlich nicht übereinstimmenden Subskalen des Fragebogens zur Selbsteinschätzung und Subskalen des Fragebogens zur Fremdeinschätzung liegen im Bereich von $.12 \leq r(304) \leq .24$ mit einem Mittelwert von $M = .16$ und p -Werten $\leq .042$ bzw. mit den nicht statistisch signifikanten Korrelationen im Bereich von $.03 \leq r(304) \leq .24$ mit einem Mittelwert von $M = .10$ und p -Werten $\leq .983$.

Korrelationen innerhalb SEC-I-SR Subskalen bzw. innerhalb SEC-I-OR Subskalen

Der Fragbogen zur Selbsteinschätzung und der Fragebogen zur Fremdeinschätzung erfassen auf ihren jeweils acht Subskalen auch verschiedene Merkmale mit der gleichen Messmethode (beispielsweise Empathisches Verhalten auf der Subskala EBE-SR im Selbstbild und die Beachtung sozialer Normen auf der Subskala SON-SR im Selbstbild). Von sämtlichen möglichen Korrelation zwischen den acht Subskalen des Fragebogens zur Selbsteinschätzung sind 96.4% (27 von 28) statistisch signifikant. Ohne die statistisch nicht signifikante Korrelation zwischen den Subskalen EBE-SR mit RES-SR ($r(304) = .08, p = .157$) liegen die Korrelationen der Subskalen des Fragebogens zur Selbsteinschätzung im Bereich von $.16 \leq r(304) \leq .75$ mit einem Mittelwert von $M = .43$ und p -Werten $\leq .004$ bzw. mit der nicht statistisch signifikanten Korrelation im Bereich von $.08 \leq r(304) \leq .75$ mit einem Mittelwert von $M = .42$ und p -Werten $\leq .157$. Von sämtlichen möglichen Korrelation

zwischen den acht Subskalen des Fragebogens zur Fremdeinschätzung sind alle (28 von 28) statistisch signifikant. Die Korrelationen der Subskalen des Fragebogens zur Fremdeinschätzung liegen im Bereich von $.27 \leq r(304) \leq .83$ mit einem Mittelwert von $M = .55$ und p -Werten $\leq .001$.

Bewertung der Multitrait-Multimethod-Analyse

Korrelationen zwischen gleichen Merkmalen, die mit verschiedenen Messmethoden erfasst wurden, fallen im Durchschnitt weitestgehend moderat aus (siehe *Korrelationen zwischen den Globalskalen SEC-SJT, SEC-SR und SEC-OR*; und *Korrelationen zwischen inhaltsgleichen SEC-I-SR und SEC-I-OR Subskalen*), jedoch deutlich höher als Korrelationen zwischen verschiedenen Merkmalen, die mit verschiedenen Messmethoden erfasst wurden (siehe *Korrelationen zwischen nicht inhaltsgleichen SEC-I-SR und SEC-I-OR Subskalen*). Damit ist die zweite Teilbedingung für das Vorliegen konvergenter und divergenter Validität erfüllt. Am höchsten fallen jedoch Korrelationen zwischen verschiedenen Merkmalen, die mit gleichen Messmethoden erfasst wurden, aus (siehe *Korrelationen zwischen der Globalskala SEC-OR und den SEC-I-SR Subskalen*; *Korrelationen zwischen der Globalskala SEC-OR und den SEC-I-OR Subskalen*; und *Korrelationen innerhalb SEC-I-SR Subskalen bzw. innerhalb SEC-I-OR Subskalen*), womit die erste Teilbedingung für das Vorliegen konvergenter und divergenter Validität nicht erfüllt ist. Es liegen Hinweise auf einen starken Einfluss der Messmethode vor. Ein erheblicher Anteil der Ähnlichkeit der mit dem Fragebogen zur Selbsteinschätzung und der mit dem Fragebogen zur Fremdeinschätzung auf den Subskalen erfassten Merkmale könnte auf die Messmethode zurückzuführen sein. Da die acht Subskalen des Fragebogens zur Selbsteinschätzung bzw. die acht Subskalen des Fragebogens zur Fremdeinschätzung jedoch sehr ähnliche, wenn auch nicht exakt die gleichen Merkmale erfassen, sind die hohen Korrelationen inhaltlich erwartungskonform. Auf dieser Basis wird empfohlen, in erster Linie die Globalskalen (SEC-SJT, SEC-SR und SEC-OR) zu interpretieren (siehe Abschnitt *Allgemeine Hinweise zur Interpretation*, S. 51). Die im Vergleich zum Fragebogen zur Selbsteinschätzung höher ausfallenden Korrelationen zwischen der Leistung im Situational Judgement Test und den Skalen des Fragebogens zur Fremdeinschätzung liefern Hinweise darauf, dass Fremdeinschätzungen etwas objektiver als Selbsteinschätzungen Aspekte des Potentials einer Testperson erfassen. Dennoch stehen auch die Skalen des Fragebogens zur Selbsteinschätzung im Zusammenhang mit der objektiven Leistung im Situational Judgement Test, was zeigt, dass auch das Selbstbild Aspekte des Potentials einer Testperson abbildet und relevante diagnostische Informationen liefert.

3.2.3.3. Divergente Validität

Allgemeine Intelligenz

Sozial-emotionale Kompetenzen sind spezifische Fähigkeiten, die durch Übung verbessert werden können (siehe Abschnitt *Theoretischer Hintergrund*, S. 18). Der entwickelte Situational Judgement Test erfasst sozial-emotionale Kompetenz als objektive Leistung einer Testperson in Form von Wissen bzw. Erfahrung (das Wissen bzw. die Erfahrung, sich sozial-emotional kompetent verhalten zu können). Der Fragebogen zur Selbsteinschätzung und der Fragebogen zur Fremdeinschätzung hingegen erfassen diese Fähigkeit(en) als subjektive Einschätzung durch die eigene (Selbstbild) bzw. eine oder mehrere andere Personen (Fremdbild). Der Erwerb von Wissen über Emotionen bzw. den Umgang mit Emotionen hängt teilweise auch von kognitiven Fähigkeiten ab, ob vorhandenes Wissen in üblichem Verhalten zum Ausdruck kommt zusätzlich von weiteren Merkmalen der Person wie auch der Situation (siehe Abschnitt *Theoretischer Hintergrund*, S. 18). Daher ist zwischen der als Leistung erfassten sozial-emotionalen Kompetenz und kognitiven Fähigkeiten ein moderater Zusammenhang zu erwarten, wohingegen Zusammenhänge zwischen subjektiven Einschätzungen und kognitiven Fähigkeiten kleiner ausfallen sollten. Zur Prüfung der Zusammenhänge zwischen den mit dem Situational Judgement Test, dem Fragebogen zur Selbsteinschätzung sowie dem Fragebogen zur Fremdeinschätzung erfassten Merkmalen und kognitiven Fähigkeiten wurde die divergente Validität zu den Standard Progressive Matrices (SPM) (R. Horn, 2009) untersucht. Die SPM sind ein sprachfreier Matrizentest mittlerer Schwierigkeit. Mit den SPM kann (fluide) allgemeine Intelligenz weitgehend unabhängig von Alter, Bildung, Nationalität und körperlicher Verfassung der Testperson erfasst werden.

Für die Ermittlung der erforderlichen Stichprobengröße zur Prüfung der Unabhängigkeit der erfassten Merkmale von (fluider) allgemeiner Intelligenz wurde eine Poweranalyse durchgeführt. Für die Poweranalyse wurde ein kleiner bis mittlerer Effekt in Höhe von $\rho = .25$ nach Cohen (1988, 1992) als relevant erachtet. Mittels G*Power (Faul et al., 2007, 2009) wurde für den Effekt von $\rho = .25$ bei einseitiger Fragestellung mit einer Irrtumswahrscheinlichkeit von $\alpha = .05$ und einer Teststärke von $1 - \beta = .80$ eine erforderliche minimale Stichprobengröße von $N = 95$ errechnet.

Dementsprechend war für die vorgesehene Prüfung divergenter Validität der entwickelten diagnostischen Verfahren bezüglich (fluider) allgemeiner Intelligenz eine Stichprobengröße von mindestens 100 Personen vorgesehen. Allerdings konnte nur eine Stichprobengröße von $n = 36$ Personen erreicht werden, bei denen zusätzlich zur vollständigen Erhebung der drei diagnostischen Verfahren auch die SPM erhoben werden konnten. Das Alter reicht in dieser Substichprobe von 16 bis 24 Jahren ($MD = 19$, $M = 18.97$, $s = 2.06$). Der Anteil an Personen weiblichen Geschlechts fällt mit 13.9% deutlich niedriger aus als der

Anteil an Personen männlichen Geschlechts mit 86.1%. Von den Personen der Substichprobe wurde in den SPM eine Leistung im Bereich von $13 \leq M \leq 56$ korrekt gelösten Matrizenaufgaben erzielt, im Mittel wurden $M = 41.67$ der insgesamt 60 Matrizenaufgaben korrekt gelöst.

Die Korrelationen der Globalskalen und Subskalen mit den SPM finden sich in Tabelle A18 des Anhangs (siehe S. 161). Es besteht ein statistisch bedeutsamer Zusammenhang zwischen der Leistung im Situational Judgement Test (Globalskala SEC-SJT) und der Leistung in den SPM, $r(34) = .33$, $p = .048$. Die Leistung im Situational Judgement Test zeigt Überschneidungen mit (fluiden) allgemeiner Intelligenz. Darüber hinaus konnten keine weiteren bedeutenden Zusammenhänge gefunden werden (was auch an der geringen Stichprobengröße und der damit einhergehenden geringen statistischen Teststärke liegen kann). Rein deskriptiv zeigt sich für die Globalskala und die Subskalen des Fragebogens zur Selbsteinschätzung ein nahezu durchgehend negativer Zusammenhang zur Leistung in den SPM, $-.26 \leq r(304) \leq .02$ mit einem Mittelwert von $M = -.13$ und p -Werten $\leq .922$. Für die Globalskala und die Subskalen des Fragebogens zur Fremdeinschätzung zeigt sich ein durchgehend positiver Zusammenhang zur Leistung in den SPM, $.03 \leq r(304) \leq .31$ mit einem Mittelwert von $M = .18$ und p -Werten $\leq .849$.

Körperliche Behinderung

In institutionellen Einrichtungen zur Förderung von Menschen mit Lernbehinderungen sind auch Menschen mit körperlicher Behinderung anzutreffen (siehe Abschnitt *Rekrutierung* der Gesamtstichprobe, S. 120; und Abschnitt *Störungsbilder und Behinderungen*, S. 121). Menschen mit körperlicher Behinderung sollten durch die speziell für die Erfassung sozial-emotionaler Kompetenz(en) bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelten Verfahren nicht diskriminiert werden (siehe Abschnitt *Normen und Subnormen*, S. 124). Daher sollten insbesondere Leistungen im Situational Judgement Test wie auch das Ergebnis einer Selbsteinschätzung durch das Vorliegen einer körperlichen Behinderung nicht negativ beeinträchtigt sein. Um eine Beeinträchtigung der Testergebnisse durch das Vorliegen einer körperlichen Behinderung zu prüfen, wurde im Rahmen der Erhebung der Gesamtstichprobe das Vorliegen einer körperlichen Behinderung dichotom erfasst und t -Tests für unabhängige Stichproben (zweiseitige Fragestellung) berechnet. Weder für den Situational Judgement Test (ohne körperliche Behinderung: $M = 6.33$, $s = 2.32$, $n = 219$; mit körperlicher Behinderung: $M = 6.82$, $s = 2.30$, $n = 79$), $t(296) = -1.63$, $p = .053$, $d = -0.21$, noch für die Skalen des Fragebogens zur Selbsteinschätzung, alle p -Werte $\geq .180$ und Effekte von $|d| \leq 0.18$, konnte ein statistisch bedeutender Einfluss gefunden werden.

Dauer der Bekanntheit

Fremdeinschätzungen sozial-emotionaler Kompetenz sollten unabhängig von der Dauer der Bekanntheit zwischen Testperson und einschätzender Person sein. Eine längere Bekanntheit könnte sowohl zu einer mildereren als auch einer strengeren Beurteilung der Testperson führen. Um einen Einfluss der Dauer der Bekanntheit auf die Skalen bzw. erfassten Merkmale des Fragebogens zur Fremdeinschätzung zu prüfen, wurde von der einschätzenden Person die Dauer der Bekanntheit mit der einzuschätzenden Person in Monaten erfragt und Pearson Korrelationen (zweiseitige Fragestellung) berechnet. Es konnte weder für die Globalskala des Fragebogens zur Fremdeinschätzung, $r(304) = .06$, $p = .319$, noch für die Subskalen des Fragebogens zur Fremdeinschätzung, $-.03 \leq r(304) \leq .10$ mit einem Mittelwert von $M = .04$ und p -Werten $\geq .092$, ein statistisch bedeutender Einfluss gefunden werden.

Bewertung der divergenten Validität

Die divergente Validität zu allgemeiner Intelligenz liefert Hinweise darauf, dass das Potential zu sozial-emotional kompetentem Verhalten mit allgemeiner Intelligenz im Zusammenhang steht. Jedoch konnte die erforderliche Stichprobengröße nicht erreicht werden, sodass eine abschließende Bewertung zum gegenwärtigen Zeitpunkt nicht möglich ist. Auch über Zusammenhänge zwischen allgemeiner Intelligenz und Selbsteinschätzungen bzw. Fremdeinschätzungen sozial-emotionaler Kompetenzen können aufgrund der zu geringen Stichprobengröße keine verlässlichen Aussagen getroffen werden. Die Prüfung der divergenten Validität an einer Stichprobe angemessener Größe (mindestens 100 Personen) steht noch aus. Die divergente Validität zu körperlicher Behinderung und Dauer der Bekanntheit zwischen eingeschätzter und einzuschätzender Person bestätigt, dass Testergebnisse durch diese beiden Faktoren nicht verzerrt werden.

3.2.3.4. Extremgruppenvalidierung

Die konkurrente Validität des Situational Judgement Tests wie auch des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung wurde für verschiedene Psychische und Verhaltensstörungen nach ICD-10 Kapitel V anhand einer Extremgruppenvalidierung geprüft. Bei einer Extremgruppenvalidierung werden Gruppen, für die a priori deutliche Unterschiede in den Ausprägungen der von einem Testverfahren zu erfassenden Merkmale zu erwarten sind, miteinander kontrastiert.

Für die Extremgruppenvalidierung wurden im Zuge der Erhebung der Gesamtstichprobe bei allen Individuen gegenwärtig diagnostizierte Störungen nach ICD-10 Kapitel V miterhoben. Gerichtete Hypothesen bezüglich zu erwartender bedeutsamer Unterschiede wurden

unabhängig von drei verschiedenen Personen mit abgeschlossener psychologischer Ausbildung, die gegenwärtig in der Forschung tätig und mit der Fachliteratur bzw. gängigen Theorien zu sozialen und emotionalen Kompetenzen vertraut sind, aufgestellt (nachfolgend als *Personen mit Expertise* bezeichnet; hierbei handelte es sich nicht um die Verfahrensentwickelnden). Für das Aufstellen der Hypothesen wurde den drei Personen mit Expertise zusätzlich die in diesem Testmanual enthaltene *Inhaltliche Beschreibung der Skalen* (siehe S. 68) zur Verfügung gestellt.

Nachfolgende Gruppen psychischer Störungen und Verhaltensstörungen wurden aufgrund zu erwartender deutlicher Unterschiede zwischen Personen ohne und Personen mit entsprechender Diagnose in den von Situational Judgement Test und Fragebogen zur Selbsteinschätzung bzw. Fragebogen zur Fremdeinschätzung erfassten Merkmale für relevant erachtet:

- ◆ Schizophrenie, schizotype und wahnhaftige Störungen (F20-F29)
- ◆ Affektive Störungen (F30-F39)
- ◆ Neurotische, Belastungs- und somatoforme Störungen (F40-F48)
- ◆ Persönlichkeits- und Verhaltensstörungen (F60-F69)
- ◆ Entwicklungsstörungen (F80-F89)
- ◆ Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend (F90-F98)

Für die genannten Störungsgruppen sind geringere Werte für die mit dem Situational Judgement Test, dem Fragebogen zur Selbsteinschätzung und dem Fragebogen zur Fremdeinschätzung erfassten Merkmale zu erwarten. Daher wurde die für einen bedeutsamen Unterschied erforderliche minimale Anzahl an Personen mit entsprechender Diagnose für eine einseitige Fragestellung für einen Effekt mittlerer Größe nach Cohen (1988, 1992) in Höhe von $d = 0.50$ mit einer Irrtumswahrscheinlichkeit von $\alpha = .05$, einer Teststärke von $1 - \beta = .80$ und einem Stichprobengrößenverhältnis von $n_a:n_b = 10.20:1.00$ bei einer Gesamtstichprobengröße von $N = 306$ mittels G*Power (Faul et al., 2007, 2009) berechnet. Um einen Effekt in Höhe von $d = 0.50$ unter diesen Bedingungen aufdecken zu können, ist jeweils eine minimale Anzahl von $n_b = 27$ Personen mit entsprechender Diagnose erforderlich. Aufgrund einer zu geringen Anzahl an Personen mit entsprechenden Diagnosen in der Gesamtstichprobe konnten die Störungsgruppe Schizophrenie, schizotype und wahnhaftige Störungen, mit einer Anzahl von $n = 5$ Personen mit entsprechender Diagnose, sowie die Störungsgruppe Persönlichkeits- und Verhaltensstörungen, mit einer Anzahl von $n = 16$ Personen mit entsprechender Diagnose, nicht in die Extremgruppenvalidierung einbezogen werden. Die Mittelwertsunterschiede für alle zum Zeitpunkt der Erhebung diagnostizierten Störungen nach ICD-10 Kapitel V können Tabelle 2 (auf der nächsten Seite) entnommen werden.

Tabelle 2

Mittelwertsunterschiede (Cohens *d*) zwischen Personen ohne Diagnose und Personen mit Diagnose für Störungsgruppen nach ICD 10 Kapitel V

Skala	F20-F29	F30-F39	F40-F48	F60-F69	F70-F79	F80-F89	F90-F98	F99
SEC-SJT	0.15	0.00	0.00	0.33	-0.23*	-0.20	-0.12	-0.02
SEC-SR	-0.11	-0.09	-0.15	-0.01	0.14	-0.15	0.27	-0.04
EBE-SR	-0.09	0.11	0.15	-0.11	-0.01	-0.15	0.32	-0.18
ECO-SR	0.36	0.12	-0.15	0.11	0.10	-0.08	0.30	-0.11
PEA-SR	0.25	-0.03	0.03	0.01	0.03	-0.15	0.40	-0.23
PES-SR	0.15	-0.25	-0.52**	-0.04	0.17	-0.14	0.02	0.27
REA-SR	-0.23	-0.03	0.14	0.05	0.07	-0.17	0.53	0.01
RES-SR	-0.40	-0.52**	-0.56**	-0.24	0.19	0.19	-0.14	0.09
SOB-SR	-0.37	0.01	0.10	0.14	0.13	-0.12	0.18	-0.20
SON-SR	-0.12	0.10	-0.12	-0.10	0.06	-0.27*	-0.17	0.18
SEC-OR	0.16	-0.26	-0.08	-0.45*	0.09	-0.31**	-0.36**	0.24
EBE-OR	0.27	-0.21	-0.02	-0.50*	0.12	-0.17	0.01	0.00
ECO-OR	-0.21	-0.01	-0.03	-0.49*	0.08	-0.37**	-0.32*	0.19
PEA-OR	-0.08	0.13	0.20	-0.34	0.07	-0.35**	-0.32**	0.23
PES-OR	0.74	0.11	-0.14	0.17	-0.07	-0.34**	-0.33*	0.46
REA-OR	-0.06	-0.41*	0.00	-0.41	0.09	-0.26*	-0.28*	0.27
RES-OR	-0.24	-0.69***	-0.54**	-0.46*	0.02	-0.12	-0.34**	0.26
SOB-OR	0.26	-0.34*	0.00	-0.45*	0.13	-0.19	-0.28*	0.10
SON-OR	0.24	-0.11	-0.06	-0.25	0.08	-0.22*	-0.40**	0.10

Anmerkung. $N = 306$. Cohens *d* mit gepoolter Standardabweichung berechnet. F20-F29 = Schizophrenie, schizotype und wahnhaftige Störungen ($n = 5$), F30-F39 = Affektive Störungen ($n = 33$), F40-F48 = Neurotische, Belastungs- und somatoforme Störungen ($n = 35$), F60-F69 = Persönlichkeits- und Verhaltensstörungen ($n = 16$), F70-F79 = Intelligenzstörung ($n = 108$), F80-F89 = Entwicklungsstörungen ($n = 96$), F90-F98 = Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend ($n = 64$), F99 = Nicht näher bezeichnete psychische Störung ($n = 51$); SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen; -SJT bzw. -SR bzw. -OR = Leistung bzw. Selbstbild bzw. Fremdbild.

* $p < .05$, einseitig. ** $p < .01$, einseitig. *** $p < .001$, einseitig.

Die Ergebnisse der Extremgruppenvalidierung werden gegliedert nach Störungsgruppen berichtet. Die Reihenfolge der Störungsgruppen orientiert sich an den von Personen mit Expertise erwarteten Unterschieden zwischen Personen ohne und Personen mit entsprechender Diagnose. Da einheitlich die meisten Unterschiede für die Störungsgruppe *Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend (F90-F98)* erwartet

wurden, werden die Ergebnisse für diese Störungsgruppe an erster Stelle berichtet. Unterschiede zwischen Personen ohne und mit entsprechender Diagnose wurden einseitig und im Falle homogener Varianzen mit dem t -Test für unabhängige Stichproben, im Fall heterogener Varianzen mit dem Welch-Test geprüft.

Störungsgruppe Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend (F90-F98)

Diese Störungsgruppe beinhaltet die Diagnosen:

- ◆ Hyperkinetische Störungen,
- ◆ Störungen des Sozialverhaltens,
- ◆ Kombinierte Störung des Sozialverhaltens und der Emotionen,
- ◆ Emotionale Störungen des Kindesalters,
- ◆ Störungen sozialer Funktionen mit Beginn in der Kindheit und Jugend sowie
- ◆ Ticstörungen (Dilling & Freyberger, 2012).

Für diese Störungsgruppe wurden einheitlich von allen Personen mit Expertise durchgehend niedrigere Werte auf allen Skalen erwartet. Für die Leistung im Situational Judgment Test konnte kein statistisch signifikanter Unterschied zwischen Personen ohne und Personen mit entsprechender Diagnose gefunden werden, $t(304) = 0.88$, $p = .380$, $d = -0.12$. Für die Skalen des Fragebogens zur Selbsteinschätzung konnten ebenfalls keine statistisch signifikanten Unterschiede zwischen Personen ohne und Personen mit entsprechender Diagnose gefunden werden, alle $ds \geq -0.17$ mit p -Werten $\geq .100$. Für die Skalen des Fragebogens zur Fremdeinschätzung konnten hingegen statistisch signifikante Unterschiede zwischen Personen ohne und Personen mit entsprechender Diagnose für die Globalskala SEC-OR (ohne entsprechende Diagnose: $M = 3.34$, $s = 0.54$, $n = 242$; mit entsprechender Diagnose: $M = 3.15$, $s = 0.54$, $n = 64$), $t(304) = 2.54$, $p = .006$, $d = -0.36$, die Subskala ECO-OR (ohne entsprechende Diagnose: $M = 3.39$, $s = 0.70$, $n = 242$; mit entsprechender Diagnose: $M = 3.17$, $s = 0.59$, $n = 64$), $t(304) = 2.30$, $p = .011$, $d = -0.32$, die Subskala PEA-OR (ohne entsprechende Diagnose: $M = 3.48$, $s = 0.62$, $n = 242$; mit entsprechender Diagnose: $M = 3.29$, $s = 0.48$, $n = 64$), $t(124,964) = 2.65$, $p = .005$, $d = -0.32$, die Subskala PES-OR (ohne entsprechende Diagnose: $M = 3.75$, $s = 0.61$, $n = 242$; mit entsprechender Diagnose: $M = 3.55$, $s = 0.61$, $n = 64$), $t(304) = 2.34$, $p = .010$, $d = -0.33$, die Subskala REA-OR (ohne entsprechende Diagnose: $M = 2.77$, $s = 0.76$, $n = 242$; mit entsprechender Diagnose: $M = 2.56$, $s = 0.80$, $n = 64$), $t(304) = 2.00$, $p = .024$, $d = -0.28$, die Subskala RES-OR (ohne entsprechende Diagnose: $M = 2.96$, $s = 0.68$, $n = 242$; mit entsprechender Diagnose: $M = 2.73$, $s = 0.75$, $n = 64$), $t(304) = 2.43$, $p = .008$, $d = -0.34$, die

Subskala SOB-OR (ohne entsprechende Diagnose: $M = 3.29$, $s = 0.65$, $n = 242$; mit entsprechender Diagnose: $M = 3.11$, $s = 0.69$, $n = 64$), $t(304) = 2.01$, $p = .023$, $d = -0.28$, und die Subskala SON-OR (ohne entsprechende Diagnose: $M = 3.88$, $s = 0.77$, $n = 242$; mit entsprechender Diagnose: $M = 3.57$, $s = 0.83$, $n = 64$), $t(304) = 2.81$, $p = .003$, $d = -0.40$, gefunden werden. Für die Subskala EBE-OR konnte kein statistisch signifikanter Unterschied zwischen Personen ohne und Personen mit entsprechender Diagnose gefunden werden, $t(304) = -0.08$, $p = .467$, $d = 0.01$.

Sowohl das Ergebnis für den Situational Judgement Test wie auch die Ergebnisse für den Fragebogen zur Selbsteinschätzung fallen damit nicht erwartungskonform aus. Für den Fragebogen zur Fremdeinschätzung fallen die Ergebnisse auf acht von neun Skalen erwartungskonform aus.

Störungsgruppe Affektive Störungen (F30-F39)

Diese Störungsgruppe beinhaltet die Diagnosen:

- ◆ Manische Episode,
- ◆ Bipolare affektive Störung,
- ◆ Depressive Episode,
- ◆ Rezidivierende depressive Störung sowie
- ◆ Anhaltende affektive Störungen (Dilling & Freyberger, 2012).

Für diese Störungsgruppe wurden einheitlich von allen Personen mit Expertise niedrigere Werte auf den Globalskalen Sozial-emotionale Kompetenz (SEC) sowie den Subskalen Wahrnehmung eigener Emotionen (PES), Regulation eigener Emotionen (RES) und Positive Beziehungsgestaltung (SOB) erwartet. Für die Leistung im Situational Judgement Test konnte kein statistisch signifikanter Unterschied zwischen Personen ohne und Personen mit entsprechender Diagnose gefunden werden, $t(304) = 0.02$, $p = .494$, $d = 0.00$. Für den Fragebogen zur Selbsteinschätzung konnte nur für die Subskala RES-SR ein statistisch signifikanter Unterschied (ohne entsprechende Diagnose: $M = 3.35$, $s = 0.88$, $n = 273$; mit entsprechender Diagnose: $M = 2.89$, $s = 0.72$, $n = 33$) gefunden werden, $t(304) = 2.83$, $p = .003$, $d = -0.52$. Auf den übrigen Skalen des Fragebogens zur Selbsteinschätzung konnten keine Unterschiede gefunden werden, alle $ds \geq -0.25$ mit p -Werten $\geq .088$. Für die Skalen des Fragebogens zur Fremdeinschätzung konnten hingegen statistisch signifikante Unterschiede zwischen Personen ohne und Personen mit entsprechender Diagnose für die Subskalen RES-OR (ohne entsprechende Diagnose: $M = 2.96$, $s = 0.67$, $n = 273$; mit entsprechender Diagnose: $M = 2.49$, $s = 0.80$, $n = 33$), $t(304) = 3.74$, $p \leq .001$, $d = -0.69$, REA-OR (ohne entsprechende Diagnose: $M = 2.76$, $s = 0.75$, $n = 273$; mit entsprechender Diagnose: $M = 2.45$, $s = 0.93$, $n = 33$), $t(304) = 2.20$, $p = .014$, $d = -0.41$, und SOB-OR

(ohne entsprechende Diagnose: $M = 3.28$, $s = 0.65$, $n = 273$; mit entsprechender Diagnose: $M = 3.05$, $s = 0.72$, $n = 33$), $t(304) = 1.84$, $p = .034$, $d = -0.34$, gefunden werden. Auf den übrigen Skalen des Fragebogens zur Fremdeinschätzung konnten keine Unterschiede gefunden werden, alle $ds \geq -0.26$ mit p -Werten $\geq .082$.

Das Ergebnis für den Situational Judgement Test fällt damit nicht erwartungskonform aus. Für den Fragebogen zur Selbsteinschätzung fallen sechs von neun Skalen erwartungskonform aus, für den Fragebogen zur Fremdeinschätzung fallen sieben von neun Skalen erwartungskonform aus.

Störungsgruppe Neurotische, Belastungs- und somatoforme Störungen (F40-F48)

Diese Störungsgruppe beinhaltet die Diagnosen:

- ◆ Phobische Störungen,
- ◆ Andere Angststörungen,
- ◆ Zwangsstörung,
- ◆ Reaktionen auf schwere Belastungen und Anpassungsstörungen,
- ◆ Dissoziative Störungen sowie
- ◆ Somatoforme Störungen (Dilling & Freyberger, 2012).

Für diese Störungsgruppe wurden einheitlich von allen Personen mit Expertise niedrigere Werte auf den Globalskalen Sozial-emotionale Kompetenz (SEC) sowie der Subskala Regulation eigener Emotionen (RES) erwartet. Für die Leistung im Situational Judgement Test konnte kein statistisch signifikanter Unterschied zwischen Personen ohne und Personen mit entsprechender Diagnose gefunden werden, $t(304) = 0.01$, $p = .992$, $d = 0.00$. Für den Fragebogen zur Selbsteinschätzung konnten statistisch signifikante Unterschiede für die Subskalen RES-SR (ohne entsprechende Diagnose: $M = 3.35$, $s = 0.87$, $n = 271$; mit entsprechender Diagnose: $M = 2.87$, $s = 0.78$, $n = 35$), $t(304) = 3.10$, $p = .001$, $d = -0.56$, und PES-SR (ohne entsprechende Diagnose: $M = 4.09$, $s = 0.75$, $n = 271$; mit entsprechender Diagnose: $M = 3.69$, $s = 0.83$, $n = 35$), $t(304) = 2.92$, $p = .002$, $d = -0.52$, gefunden werden. Auf den übrigen Skalen des Fragebogens zur Selbsteinschätzung konnten keine Unterschiede gefunden werden, alle $ds \geq -0.15$ mit p -Werten $\geq .203$. Für den Fragebogen zur Selbsteinschätzung konnte nur für die Subskala RES-OR (ohne entsprechende Diagnose: $M = 2.96$, $s = 0.66$, $n = 271$; mit entsprechender Diagnose: $M = 2.59$, $s = 0.86$, $n = 35$) ein statistisch signifikanter Unterschied gefunden werden, $t(39,394) = 2.45$, $p = .010$, $d = -0.54$. Auf den übrigen Skalen der Fremdeinschätzung konnten keine Unterschiede gefunden werden, alle $ds \geq -0.14$ mit p -Werten $\geq .215$.

Das Ergebnis für den Situational Judgement Test fällt damit nicht erwartungskonform aus. Für den Fragebogen zur Selbsteinschätzung fallen sieben von neun Skalen erwartungskonform aus, für den Fragebogen zur Fremdeinschätzung fallen acht von neun Skalen erwartungskonform aus.

Störungsgruppe Entwicklungsstörungen (F80-F89)

Diese Störungsgruppe beinhaltet die Diagnosen:

- ◆ Umschriebene Entwicklungsstörungen des Sprechens und der Sprache,
- ◆ Umschriebene Entwicklungsstörungen schulischer Fertigkeiten,
- ◆ Umschriebene Entwicklungsstörung der motorischen Funktionen,
- ◆ Kombinierte umschriebene Entwicklungsstörungen sowie
- ◆ Tief greifende Entwicklungsstörungen (Dilling & Freyberger, 2012).

Für diese Störungsgruppe wurde von den Personen mit Expertise keine bzw. eine nicht einheitliche Beeinträchtigung der Ausprägungen auf den Skalen erwartet. Für die Leistung im Situational Judgement Test konnte kein statistisch signifikanter Unterschied zwischen Personen ohne und Personen mit entsprechender Diagnose gefunden werden, $t(304) = 1.61$, $p = .054$, $d = -0.20$. Für den Fragebogen zur Selbsteinschätzung konnte für die Subskala SON-SR (ohne entsprechende Diagnose: $M = 4.37$, $s = 0.58$, $n = 210$; mit entsprechender Diagnose: $M = 4.21$, $s = 0.68$, $n = 96$) ein statistisch signifikanter Unterschied gefunden werden, $t(304) = 2.16$, $p = .016$, $d = -0.27$. Auf den übrigen Skalen des Fragebogens zur Selbsteinschätzung konnten keine Unterschiede gefunden werden, alle $ds \geq -0.17$ mit p -Werten $\geq .080$. Für den Fragebogen zur Fremdeinschätzung konnten statistisch signifikante Unterschiede für die Globalskala SEC-OR (ohne entsprechende Diagnose: $M = 3.35$, $s = 0.51$, $n = 210$; mit entsprechender Diagnose: $M = 3.18$, $s = 0.61$, $n = 96$), $t(304) = 2.55$, $p = .006$, $d = -0.31$, sowie für die Subskalen ECO-OR (ohne entsprechende Diagnose: $M = 3.42$, $s = 0.63$, $n = 210$; mit entsprechender Diagnose: $M = 3.17$, $s = 0.75$, $n = 96$), $t(158,442) = 2.78$, $p = .003$, $d = -0.37$, PEA-OR (ohne entsprechende Diagnose: $M = 3.51$, $s = 0.56$, $n = 210$; mit entsprechender Diagnose: $M = 3.30$, $s = 0.65$, $n = 96$), $t(304) = 2.81$, $p = .003$, $d = -0.35$, PES-OR (ohne entsprechende Diagnose: $M = 3.78$, $s = 0.55$, $n = 210$; mit entsprechender Diagnose: $M = 3.57$, $s = 0.71$, $n = 96$), $t(149,337) = 2.53$, $p = .006$, $d = -0.34$, REA-OR (ohne entsprechende Diagnose: $M = 2.79$, $s = 0.73$, $n = 210$; mit entsprechender Diagnose: $M = 2.59$, $s = 0.85$, $n = 96$), $t(304) = 2.08$, $p = .019$, $d = -0.26$, und SON-OR (ohne entsprechende Diagnose: $M = 3.87$, $s = 0.75$, $n = 210$; mit entsprechender Diagnose: $M = 3.69$, $s = 0.87$, $n = 96$), $t(304) = 1.81$, $p = .036$, $d = 0.22$, gefunden werden. Auf den übrigen Skalen des Fragebogens zur Fremdeinschätzung konnten keine Unterschiede gefunden werden, alle $ds \geq -0.19$ mit p -Werten $\geq .066$.

Die Störungsgruppe Entwicklungsstörungen ist deutlich heterogener als andere Störungsgruppen und umfasst neben der Störung der Entwicklung verschiedener spezifischer Fertigkeiten auch Störungen aus dem Autismus-Spektrum. Hierdurch ist eine gezielte Vorhersage einer Beeinträchtigung der Ausprägungen auf den Skalen erschwert. Für den Situational Judgement Test wie auch die meisten Skalen des Fragebogens zur Selbsteinschätzung konnten keine Unterschiede gefunden werden. Für den Fragebogen zur Fremdeinschätzung finden sich Unterschiede auf sechs von neun Skalen, die weitestgehend auf Störungen aus dem Autismus-Spektrum zurückzuführen sein dürften. Eine weitere Prüfung der konkurrenten Validität auf der Basis differenzierterer Diagnosen aus dem Bereich der Entwicklungsstörungen sollte noch erfolgen.

Störungsgruppe Intelligenzstörung (F70-F79)

Diese Störungsgruppe beinhaltet die Diagnosen:

- ◆ Leichte Intelligenzminderung,
- ◆ Mittelgradige Intelligenzminderung,
- ◆ Schwere Intelligenzminderung,
- ◆ Schwerste Intelligenzminderung sowie
- ◆ Dissoziierte Intelligenz (Dilling & Freyberger, 2012).

Menschen mit geistiger Behinderung sollten durch die speziell für die Erfassung sozial-emotionaler Kompetenz(en) bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelten Verfahren nicht diskriminiert werden (siehe Abschnitt *Normen und Subnormen*, S. 124). Daher sollten Leistungen im Situational Judgement Test wie auch Ergebnisse aus dem Fragebogen zur Selbsteinschätzung und dem Fragebogen zur Fremdeinschätzung durch das Vorliegen einer geistigen Behinderung nicht negativ beeinträchtigt sein. Für die Leistung im Situational Judgement Test konnte ein statistisch signifikanter Unterschied (ohne entsprechende Diagnose: $M = 6.65$, $s = 2.34$, $n = 198$; mit entsprechender Diagnose: $M = 6.11$, $s = 2.24$, $n = 108$) gefunden werden, $t(304) = 1.96$, $p = .026$, $d = -0.23$. Auf den Skalen des Fragebogens zur Selbsteinschätzung, alle $ds \geq -0.01$ mit p -Werten $\geq .473$, und den Skalen des Fragebogens zur Fremdeinschätzung, alle $ds \geq -0.07$ mit p -Werten $\geq .290$, konnten keine Unterschiede gefunden werden.

Das Ergebnis für den Situational Judgement Test fällt damit nicht erwartungskonform aus. Für den Fragebogen zur Selbsteinschätzung fallen neun von neun Skalen erwartungskonform aus, für den Fragebogen zur Fremdeinschätzung fallen ebenfalls neun von neun Skalen erwartungskonform aus.

Bewertung der Extremgruppenvalidierung

Die Leistung im Situational Judgement Test scheint von den verschiedenen hier geprüften Störungen weitestgehend unabhängig zu sein und am stärksten durch eine Störung kognitiver Fähigkeiten beeinträchtigt zu werden. Die Fremdeinschätzung fällt in höherem Maße als die Selbsteinschätzung erwartungskonform für verschiedene Störungsgruppen aus.

3.2.3.5. Konkurrente Validität

Bildungsstand

Meta-analytische Befunde (MacCann et al., 2019; Sánchez-Álvarez et al., 2020) lassen Zusammenhänge zwischen sozial-emotionalen Kompetenzen und schulischen Leistungen erwarten, wobei kognitive Aspekte von wesentlicher Bedeutung sind. Um Zusammenhänge zwischen der Skala des Situational Judgement Tests, den Skalen des Fragebogens zur Selbsteinschätzung sowie den Skalen des Fragebogens zur Fremdeinschätzung und schulischer Leistung zu testen, wurde der Bildungsstand in Form des Schulabschlusses (Förderschul-, Hauptschul-/Mittelschul-, Realschul-, Fachoberschul-/Gymnasialabschluss) miterfasst und Pearson Korrelationen (zweiseitige Fragestellung) berechnet. Die Ergebnisse können Tabelle 3 (auf der nächsten Seite) entnommen werden. Sowohl die Globalskala des Situational Judgement Tests, $r(304) = .17, p = .003$, wie auch die meisten Skalen des Fragebogens zur Fremdeinschätzung, $.13 \leq r(304) \leq .19$ mit einem Mittelwert von $M = .16$ und p -Werten $\leq .030$ (bis auf die Subskala EBE-OR, $r(304) = .09, p = .139$, und die Subskala REA-OR, $r(304) = .09, p = .108$), stehen im Zusammenhang mit dem Bildungsstand. Entgegen der Erwartung konnte für die Skalen des Fragebogens zur Selbsteinschätzung kein Zusammenhang mit dem Bildungsstand gefunden werden, $-.08 \leq r(304) \leq .10$ mit einem Mittelwert von $M = .02$ und p -Werten $\geq .080$.

Ausbildungsstatus

Auch für die berufliche Leistung lassen sich aufgrund meta-analytischer Befunde (Joseph et al., 2015; O'Boyle et al., 2011) Zusammenhänge mit sozial-emotionalen Kompetenzen erwarten. Um Zusammenhänge zwischen der Skala des Situational Judgement Tests, den Skalen des Fragebogens zur Selbsteinschätzung sowie den Skalen des Fragebogens zur Fremdeinschätzung und beruflicher Leistung zu testen, wurde der Ausbildungsstatus (nicht in Ausbildung, in Ausbildung) dichotom miterfasst und Pearson Korrelationen berechnet. Die Ergebnisse können Tabelle 3 (auf der nächsten Seite) entnommen werden. Sowohl die Skala des Situational Judgement Tests, $r(304) = .13, p = .026$, wie auch die meisten Skalen des Fragebogens zur Fremdeinschätzung, $.12 \leq r(304) \leq .26$ mit einem Mittelwert von $M = .17$ und p -Werten $\leq .040$ (bis auf die Subskala PES-OR, $r(304) = .10, p = .097$), stehen

im Zusammenhang mit dem Ausbildungsstatus. Entgegen der Erwartung konnte für die Skalen des Fragebogens zur Selbsteinschätzung kein Zusammenhang mit dem Ausbildungsstatus gefunden werden, $-.01 \leq r(304) \leq .10$ mit einem Mittelwert von $M = .05$ und p -Werten $\geq .091$.

Tabelle 3

Pearson Korrelationen zwischen den Skalen des Situational Judgement Tests (SEC-SJT), des Fragebogens zur Selbsteinschätzung (SEC-I-SR) sowie des Fragebogens zur Fremdeinschätzung (SEC-I-OR) und Bildungsstand sowie Ausbildungsstatus

Skala	Bildungsstand	Ausbildungsstatus
SEC-SJT	0,17**	0,13*
SEC-SR	0.02	0.06
EBE-SR	-0.02	0.02
ECO-SR	0.06	0.03
PEA-SR	0.03	0.07
PES-SR	0,10	0,10
REA-SR	-0.08	-0.01
RES-SR	-0.01	0.02
SOB-SR	-0.02	0.05
SON-SR	0.07	0.08
SEC-OR	0,18**	0,20**
EBE-OR	0.09	0,12*
ECO-OR	0,14*	0,19**
PEA-OR	0,16**	0,20**
PES-OR	0,17**	0,10
REA-OR	0.09	0,26**
RES-OR	0,13*	0,17**
SOB-OR	0,16**	0,12*
SON-OR	0,19**	0,13*

Anmerkung. $N = 306$. Zur besseren Vergleichbarkeit sind Mittelwertsunterschiede im Ausbildungsstatus auch als Pearson Korrelation angegeben. SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen; -SJT bzw. -SR bzw. -OR = Leistung bzw. Selbstbild bzw. Fremdbild.

* $p < 0.05$. ** $p < 0.01$.

Bewertung der konkurrenten Validität

Sowohl der Situational Judgement Test wie auch der Fragebogen zur Fremdeinschätzung weisen konkurrente Validität zu Bildungsstand und Ausbildungsstatus auf. Für den Fragebogen zur Selbsteinschätzung konnten für die beiden zusätzlichen Außenkriterien keine Hinweise auf konkurrente Validität gefunden werden.

3.2.3.6. Gesamtbewertung der Validität

Die Ergebnisse der Validierung liefern Hinweise darauf, dass die von den jeweiligen Verfahren erfassten sozial-emotionalen Kompetenzen relativ homogen ausfallen (siehe Abschnitt *Faktorielle Validität*, S. 97; und Abschnitt *Multitrait-Multimethod-Analyse*, S. 101). Die Leistung im Situational Judgement Test, die Selbsteinschätzung und die Fremdeinschätzung scheinen verschiedene, jedoch nicht voneinander unabhängige Aspekte sozial-emotionaler Kompetenz abzubilden (siehe Abschnitt *Multitrait-Multimethod-Analyse*, S. 101). Das Potential zu sozial-emotional kompetentem Verhalten (Leistung im Situational Judgement Test) scheint am deutlichsten mit kognitiven Fähigkeiten in Zusammenhang zu stehen, wie die Befunde zu allgemeiner Intelligenz, Intelligenzstörungen, Bildungsstand und Ausbildungsstatus zeigen (siehe Abschnitt *Divergente Validität*, S. 106; Abschnitt *Extremgruppenvalidierung*, S. 108; und Abschnitt *Konkurrente Validität*, S. 116). Die Befunde für den Fragebogen zur Fremdeinschätzung deuten darauf hin, dass auch übliches sozial-emotional kompetentes Verhalten mit kognitiven Fähigkeiten im Zusammenhang steht, wie die Befunde zu Bildungsstand und Ausbildungsstatus zeigen (siehe Abschnitt *Konkurrente Validität*, S. 116). Im Gegensatz zum Fragebogen zur Fremdeinschätzung scheint der Fragebogen zur Selbsteinschätzung für sämtliche Außenkriterien nur über eine stark eingeschränkte Validität zu verfügen, was auf eine mangelnde Fähigkeit zur Selbstreflexion bzw. Selbstüberschätzung oder -unterschätzung hindeuten könnte. Die Ergebnisse aus dem Fragebogen zur Selbsteinschätzung sollten dementsprechend vorsichtig interpretiert werden.

4. Normierung

Das vierte Kapitel widmet sich einer Beschreibung der Rekrutierung sowie demografischer Merkmale der für die Normierung herangezogenen Gesamtstichprobe. Es wird auf die Notwendigkeit von Normen für Substichproben weiblichen und männlichen Geschlechts eingegangen. Die Normierung wird ausführlich beschrieben.

4.1. Rekrutierung der Gesamtstichprobe

Für die Normierung war ursprünglich eine Gesamtstichprobengröße von mindestens 500 Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung vorgesehen. Im Zeitraum von Januar bis März 2020 konnten jedoch nur 399 Individuen erhoben werden. Sämtliche Daten wurden in institutionellen Einrichtungen zur Förderung von Menschen mit Lernbehinderungen erhoben. Dort nehmen die Jugendlichen und jungen Erwachsenen an einer Berufsvorbereitenden Bildungsmaßnahme oder dem Vorqualifizierungsjahr Arbeit und Beruf teil.

Die Jugendlichen und jungen Erwachsenen wurden im Vorfeld der Erhebung um ihre Unterstützung bei der Entwicklung von psychologischen Tests zur Erfassung sozial-emotionaler Kompetenzen gebeten. Alle Jugendlichen und jungen Erwachsenen konnten freiwillig an der Normierung teilnehmen. Ihnen wurde versichert, dass jeder Beitrag wertvoll ist. Alle Daten wurden nur bei vorliegendem Einverständnis der jungen Erwachsenen bzw. eines gesetzlichen Vertreters erhoben.

Für die Mitarbeit an der Erhebung danken wir allen Beteiligten des Bundesverbands LERNEN FÖRDERN e.V. (Remseck), des Berufsbildungswerks Josefsheim (Bigge), des Berufsbildungswerks Mosbach-Heidelberg (Mosbach und Heidelberg), des Berufsbildungswerks Benediktushof Maria Veen (Reken), der Mariaberger Ausbildung & Service gemeinnützige GmbH (Gammertingen), der Bodelschwingschen Stiftungen Bethel (Bielefeld) und des staatlich regionalen Förderzentrums Rothenbachschule (Sömmerda).

4.2. Beschreibung der Gesamtstichprobe

Das Inventar und der Situational Judgement Test wurden speziell für die multimethodale Erfassung sozial-emotionaler Kompetenz(en) bei Jugendlichen und jungen Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelt und entsprechend normiert. Die Stichprobe erhebt daher für kein soziodemografisches Merkmal (Durchschnittsalter, Geschlechterverhältnis, Bildungsniveau) Anspruch auf Repräsentativität für die deutsche Bevölkerung. Vielmehr handelt es sich um eine anfallende Stichprobe von Individuen aus institutionellen Einrichtungen zur Förderung von Menschen mit Lernbehinderungen, deren Repräsentativität auch für diesen Teil der Bevölkerung eingeschränkt sein könnte.

4.2.1. Alter und Geschlecht

Aufgrund von fehlenden Angaben für Alter und Geschlecht, ungeeignetem Alter (jünger als 14 und älter als 24 Jahre), einer zu geringen Anzahl von Individuen in einer Altersklasse (weniger als 10) und fehlender Testwerte war es erforderlich 93 Individuen auszuschließen. Es konnten Daten von $N = 306$ Jugendlichen und jungen Erwachsenen im Alter von 15 bis 24 Jahren ($MD = 17$, $M = 18.72$, $s = 2.32$) in die Gesamtstichprobe einfließen. Bei der Zusammensetzung der Gesamtstichprobe für die Normierung wurde darauf geachtet, dass jede der Altersklassen mit mindestens zehn Individuen besetzt ist. Auf die am stärksten besetzte Altersklasse von 17 Jahren entfielen 56 Individuen (18.3%), auf die am geringsten besetzte Altersklasse von 24 Jahren 12 Individuen (3.9%). Der Anteil an Individuen weiblichen Geschlechts fällt mit 36.3% geringer aus als der Anteil an Individuen männlichen Geschlechts (111 weiblich; 195 männlich, 63.7%). Hinsichtlich des Alters besteht in der Gesamtstichprobe zwischen den beiden Geschlechtern weiblich ($M = 18.86$, $s = 2.52$) und männlich ($M = 18.64$, $s = 2.21$) kein statistisch bedeutsamer Unterschied, $t(304) = 0.83$, $p = .408$, $d = 0.09$.

4.2.2. Bildungsstand und Ausbildungsstatus

Ein Großteil der Jugendlichen und jungen Erwachsenen war im Besitz eines Förderschul- (139 Individuen, 45.4%) oder eines Hauptschul-/Mittelschulabschlusses (95 Individuen, 31.0%), ein geringer Teil im Besitz eines Realschul- (18 Individuen, 5.9%) oder eines Fachoberschul-/Gymnasialabschlusses (12 Individuen, 3.9%). 33 Individuen (10.8%) waren ohne Schulabschluss. Insgesamt befanden sich zum Zeitpunkt der Erhebung 169 Jugendliche und junge Erwachsene (55.2%) in Ausbildung.

4.2.3. Störungsbilder und Behinderungen

Bereits im Vorfeld der Datenerhebung waren bei den 306 Jugendlichen und jungen Erwachsenen folgende Störungsbilder diagnostiziert worden (Mehrfachdiagnosen können vorliegen; Nennung nach Häufigkeit in absteigender Reihenfolge): Intelligenzstörung (108 Individuen, 35.3%); Entwicklungsstörungen (96 Individuen, 31.4%); Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend (64 Individuen, 20.9%); Nicht näher bezeichnete psychische Störungen (51 Individuen, 16.7%); Neurotische, Belastungs- und somatoforme Störungen (35 Individuen, 11.4%); Affektive Störungen (33 Individuen, 10.8%); Persönlichkeits- und Verhaltensstörungen (16 Individuen, 5.2%); Schizophrenie, schizotype und wahnhaftige Störungen (5 Individuen, 1.6%).

Menschen mit Behinderung sind nach §2 SGB IX und §3 BGG Menschen, die eine länger als sechs Monate andauernde körperliche, seelische, geistige oder Sinnesbeeinträchtigung an der gleichberechtigten Teilhabe an der Gesellschaft hindert. Bei 182 Individuen (59.5%) lag zum Zeitpunkt der Erhebung ein Grad der Behinderung von 0 vor, bei 46 Individuen (15.0%) ein Grad der Behinderung von mindestens 20 und weniger als 50, bei 78 Individuen (25.5%) ein Grad der Behinderung von mindestens 50. Bei 79 Individuen (25.8%) lag auch eine körperliche Beeinträchtigung vor.

4.3. Berechnung der Normwerte

4.3.1. Verteilung der Skalenrohwerte

Normal verteilte Skalenrohwerte können mittels linearer Transformation in standardisierte Werte überführt werden, anormal verteilte Skalenrohwerte müssen zuvor normalisiert werden. Zur Beurteilung einer Verteilung von Werten kann deren Histogramm einer visuellen Inspektion unterzogen werden. Neben der visuellen Inspektion der Verteilung kann auch deren Schiefe und Exzess für die Beurteilung der Annäherung an die Normalverteilung herangezogen werden. Die Schiefe beschreibt die Symmetrie einer Verteilung, der Exzess die Steilheit einer Verteilung in Relation zur Normalverteilung ($\text{Exzess} = \text{Kurtosis} - 3$). Die Normalverteilung weist eine Schiefe bzw. einen Exzess von 0 auf. Eine Schiefe bzw. ein Exzess kleiner 0 weisen auf eine linksschiefe bzw. flachere Verteilung hin, eine Schiefe bzw. ein Exzess größer als 0 auf eine rechtsschiefe bzw. steilere Verteilung. Inferenzstatistische Tests sind zur Prüfung der Normalverteilung nur bedingt geeignet, da mit zunehmender Stichprobengröße schon geringfügige Abweichungen von der Normalverteilung statistisch bedeutsam werden.

Das Histogramm der Globalskala Sozial-emotionale Kompetenz des Situational Judgement Tests (SEC-SJT) (siehe Abbildung A1, S. 162) zeigt eine deutlich linksschiefe Verteilung. Auch die Histogramme der Subskala Wahrnehmung eigener Emotionen (PES-SR) und der Subskala Beachtung sozialer Normen (SON-SR) des Fragebogens zur Selbsteinschätzung (siehe Abbildung A2, S. 163) und der Subskala Beachtung sozialer Normen des Fragebogens zur Fremdeinschätzung (SON-OR) (siehe Abbildung A3, S. 164) zeigen deutlich linksschiefe Verteilungen.

Mittelwerte, Standardabweichungen, Schiefen und Exzesse der Skalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung sind für die Gesamtstichprobe ($N = 306$) in Tabelle A17 des Anhangs (siehe S. 160) angegeben. Bei den Skalenkennwerten zeigen sowohl die Globalskala des Situational Judgement Tests wie auch

die Globalskalen und Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung in Schiefen und Exzessen (geringfügige) Abweichungen von der Normalverteilung. Bis auf die bereits bei der visuellen Inspektion auffälligen Skalen fallen Schiefen und Exzesse der übrigen Skalen ihrem Betrag nach jedoch zumeist kleiner als 0.50 aus. Alle Skalen zeigen eine Tendenz zur Linksschiefe. Die Skalen des Fragebogens zur Fremdeinschätzung zeigen im Vergleich zu der Globalskala des Situational Judgement Tests und den Skalen des Fragebogens zur Selbsteinschätzung fast durchgehend kleinere Abweichungen von der Normalverteilung.

4.3.2. Normierung der Rohwerte

4.3.2.1. Normierung des Situational Judgement Tests

Aufgrund relativ anormal verteilter Rohwerte und der durch die niedrige Itemanzahl bedingten geringen Variationsbreite der Rohwerte wurden den Rohwerten des Situational Judgement Tests nur die entsprechenden Prozentränge der Stichprobe zugeordnet. Für den Situational Judgement Test liegt daher nur eine Prozentrangnorm vor.

4.3.2.2. Normierung des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung

Auch wenn die meisten der Skalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung keine gravierenden Abweichungen von der Normalverteilung zeigen, sind doch für alle Skalen die Rohwerte zumindest geringfügig anormal verteilt. Um hierauf adäquat einzugehen und anormal verteilte Skalenrohwerte zu normalisieren wurden die Skalenrohwerte der Globalskalen und der Subskalen des Fragebogens zur Selbsteinschätzung und des Fragebogens zur Fremdeinschätzung durchgehend mit der nichtlinearen Flächentransformation nach McCall (1922) in T -Werte überführt. Die derart berechneten T -Werte können als normalisierte standardisierte Werte einer anormalen Rohwertverteilung verstanden werden. Aufgrund der relativ kleinen Gesamtstichprobe und der durchgehenden Tendenz zur Linksschiefe der Skalen sind nicht alle theoretisch realisierbaren Skalenrohwerte in der Gesamtstichprobe vorhanden, wodurch sich Sprünge in der Reihe der Rohwerte befinden (siehe Abschnitt *Transformation von Rohwerten in standardisierte Werte*, S. 59). Jedem T -Wert kann auch ein bestimmter Prozentrang zugeordnet werden. Für die Skalen des Fragebogens zur Selbsteinschätzung und die Skalen des Fragebogens zur Fremdeinschätzung liegen daher T -Wert-Normen und Prozentrangnormen vor.

4.3.3. Normen und Subnormen

Ein psychologischer Test gilt als fair, wenn keine der mit dem Test untersuchbaren Gruppen diskriminiert wird. Finden sich in den von einem Test erfassten Merkmalen praktisch bedeutsame Unterschiede für ein Gruppierungsmerkmal, das für die erfassten Merkmale (theoretisch) nicht von Bedeutung sein sollte, gilt dieser Test als potentiell unfair. Um dieser potentiellen Diskriminierung von Personen verschiedener Gruppenzugehörigkeit entgegenwirken zu können, werden separate (Sub-)Normen für die betroffenen Gruppen zur Verfügung gestellt. Für den Situational Judgement Test sowie den Fragebogen zur Selbsteinschätzung und den Fragebogen zur Fremdeinschätzung wurden Gruppenunterschiede für das soziodemografische Merkmal Geschlecht mit dem t -Test für unabhängige Stichproben (zweiseitige Fragestellung) geprüft. Mittelwerte, Standardabweichungen und Cohens d s sind in Tabelle 4 (auf der nächsten Seite) angegeben.

Es wurden statistisch bedeutsame Unterschiede auf den folgenden Skalen gefunden:

- ◆ der Globalskala SEC-SJT (weiblich: $M = 6.95$, $s = 2.17$, $n = 111$; männlich: $M = 6.18$, $s = 2.36$, $n = 195$), $t(304) = 2.80$, $p = .003$, $d = -0.33$;
- ◆ der Subskala EBE-SR (weiblich: $M = 3.90$, $s = 0.75$, $n = 111$; männlich: $M = 3.70$, $s = 0.79$, $n = 195$), $t(304) = 2.17$, $p = .015$, $d = -0.26$;
- ◆ der Subskala RES-SR (weiblich: $M = 3.01$, $s = 0.88$, $n = 111$; männlich: $M = 3.46$, $s = 0.83$, $n = 195$), $t(304) = -4.47$, $p \leq .001$, $d = 0.53$;
- ◆ der Subskala SOB-SR (weiblich: $M = 3.79$, $s = 0.56$, $n = 111$; männlich: $M = 3.59$, $s = 0.62$, $n = 195$), $t(304) = 2.74$, $p = .003$, $d = -0.33$;
- ◆ der Subskala EBE-OR (weiblich: $M = 3.39$, $s = 0.78$, $n = 111$; männlich: $M = 3.16$, $s = 0.78$, $n = 195$), $t(304) = 2.49$, $p = .007$, $d = -0.30$; und
- ◆ der Subskala RES-OR (weiblich: $M = 2.73$, $s = 0.72$, $n = 111$; männlich: $M = 3.02$, $s = 0.66$, $n = 195$), $t(304) = -3.45$, $p = .001$, $d = 0.41$.

Da sowohl für den Situational Judgement Test wie auch den Fragebogen zur Selbsteinschätzung und den Fragebogen zur Fremdeinschätzung statistisch bedeutsame Unterschiede gefunden werden konnten, werden für alle drei Verfahren zusätzlich zur Norm für die Gesamtstichprobe separate Normen für Substichproben weiblichen und männlichen Geschlechts zur Verfügung gestellt.

Berechnung der Normwerte

Tabelle 4

Mittelwerte, Standardabweichungen und Cohens *d*s für Substichproben weiblichen und männlichen Geschlechts

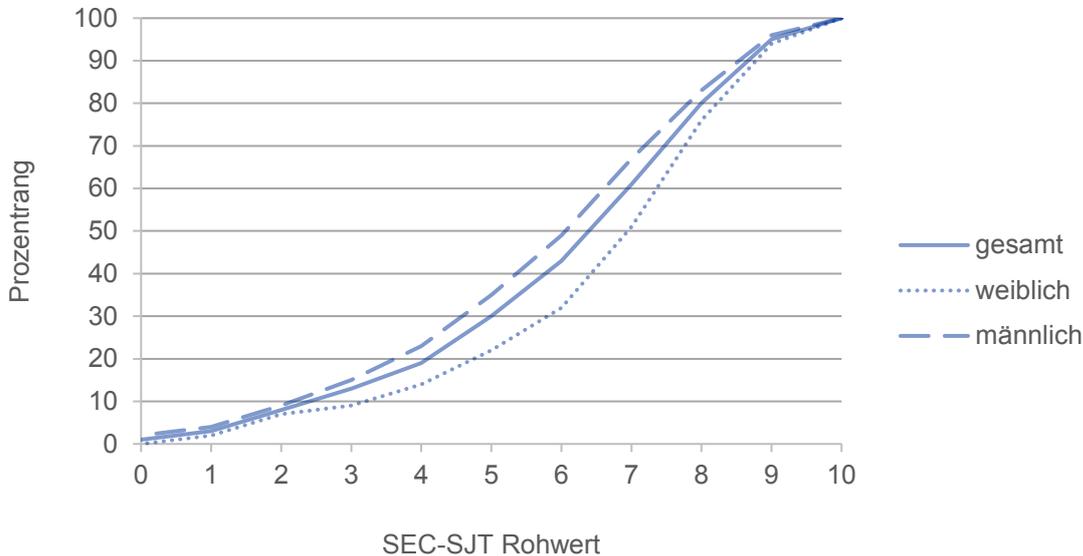
Skala	<i>M</i> (<i>s</i>)		Cohens <i>d</i>
	weiblich (<i>n</i> = 111)	männlich (<i>n</i> = 195)	
SEC-SJT	6.95 (2.17)	6.18 (2.36)	-0.33**
SEC-SR	3.70 (0.50)	3.67 (0.51)	-0.05
EBE-SR	3.90 (0.75)	3.70 (0.79)	-0.26*
ECO-SR	3.42 (0.68)	3.51 (0.69)	0.14
PEA-SR	3.67 (0.64)	3.67 (0.71)	0.00
PES-SR	3.99 (0.81)	4.08 (0.74)	0.11
REA-SR	3.37 (0.75)	3.24 (0.75)	-0.17
RES-SR	3.01 (0.88)	3.46 (0.83)	0.53**
SOB-SR	3.79 (0.56)	3.59 (0.62)	-0.33**
SON-SR	4.40 (0.58)	4.28 (0.63)	-0.21
SEC-OR	3.34 (0.56)	3.28 (0.54)	-0.11
EBE-OR	3.39 (0.78)	3.16 (0.78)	-0.30*
ECO-OR	3.35 (0.67)	3.34 (0.69)	-0.02
PEA-OR	3.49 (0.58)	3.42 (0.61)	-0.11
PES-OR	3.76 (0.58)	3.68 (0.63)	-0.13
REA-OR	2.80 (0.81)	2.69 (0.76)	-0.14
RES-OR	2.73 (0.72)	3.02 (0.66)	0.41**
SOB-OR	3.34 (0.69)	3.21 (0.64)	-0.20
SON-OR	3.82 (0.87)	3.81 (0.74)	-0.01

Anmerkung. Cohens *d* mit gepoolter Standardabweichung berechnet. *M* = Mittelwert; *s* = Standardabweichung; SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen; -SJT bzw. -SR bzw. -OR = Leistung bzw. Selbstbild bzw. Fremdbild.

Abbildung 20 (auf der nächsten Seite) veranschaulicht den Verlauf der kumulierten Häufigkeit von Personen (Prozentrang) in Abhängigkeit vom Rohwert der Globalskala des Situational Judgement Tests (SEC-SJT) und der Stichprobe. Aufgrund der durchschnittlich besseren Leistung weiblicher Personen (in der Gesamtstichprobe beantworteten weibliche Personen durchschnittlich fast ein Item mehr korrekt) erreichen männliche Testpersonen bei gleichem Rohwert einen höheren Prozentrang als weibliche Testpersonen. Wird für die Ermittlung des Normwerts die Gesamtstichprobe herangezogen, wird dieser Effekt ausgeglichen.

Abbildung 20

Prozentrang in Abhängigkeit vom Rohwert der Globalskala des Situational Judgement Tests (SEC-SJT) und der Stichprobe



Anmerkung. $N_{\text{gesamt}} = 306$, $n_{\text{weiblich}} = 111$, $n_{\text{männlich}} = 195$.

Im Fragebogen zur Selbsteinschätzung schätzen weibliche Testpersonen ihre Bereitschaft, andere Personen zu unterstützen (Subskala EBE-SR) und ihr Bemühen um den Aufbau und die Aufrechterhaltung positiver Beziehungen zu anderen (Subskala SOB-SR) höher und ihre Fähigkeit zur Regulation eigener Emotionen (Subskala RES-SR) niedriger ein als männliche Testpersonen. Die Effekte im Fragebogen zur Fremdeinschätzung weisen in die gleiche Richtung, auf der Subskala Empathisches Verhalten (EBE-OR) werden weibliche Testpersonen höher und auf der Subskala Regulation eigener Emotionen (RES-OR) niedriger eingeschätzt als männliche Testpersonen.

Welche Norm für die Interpretation der Testergebnisse einer Testperson am geeignetsten ist, hängt weitestgehend von der diagnostischen Fragestellung ab. Schätzt beispielsweise eine weibliche Testperson ihre Fähigkeit, regulierend auf das eigene emotionale Erleben einwirken und unangenehme Emotionen bewältigen zu können (Subskala RES-SR), selbst ähnlich durchschnittlich wie eine männliche Testperson ein, wird die weibliche Testperson bei Wahl der Gesamtstichprobe ein niedrigeres Testergebnis als die männliche Testperson erreichen (denn durchschnittlich schätzen sich weibliche Personen auf der Subskala RES-SR selbst niedriger ein). Bei Wahl der Substichprobe weiblichen Geschlechts wird diese weibliche Testperson hingegen einen ähnlichen T -Wert wie die männliche Testperson in der Substichprobe männlichen Geschlechts erreichen. Die Wahl der geschlechtsspezifischen Substichprobe gleicht den Effekt aus, dass sich weibliche Personen durchschnittlich niedriger einschätzen als männliche Personen. Ob solch ein Ausgleich angemessen ist,

zeigt sich bei Einbezug der diagnostischen Fragestellung: Sollen beispielsweise förderungsbedürftige Personen identifiziert werden, ist es durchaus möglich, dass eine weibliche Testperson auf der Subskala RES-SR bei Wahl der weiblichen Substichprobe keinen Bedarf für Förderung zeigt, wohingegen sie bei Wahl der Gesamtstichprobe Bedarf für Förderung zeigen könnte (weil es weiblichen Personen ihrer eigenen Einschätzung nach schwerer fällt als männlichen Personen, regulierend auf das eigene emotionale Erleben einzuwirken und unangenehme Emotionen zu bewältigen). Bei der Wahl einer angemessenen Norm handelt es sich um eine Entscheidung der Interpretation, der kein allgemeingültiges Schema zugrunde gelegt werden kann. Daher sollte diese Entscheidung auch möglichst nur von entsprechend ausgebildetem Fachpersonal getroffen werden.

4.3.3.1. Normwerte für die Skalen des Fragebogens zur Selbsteinschätzung (SEC-I-SR)

T-Werte und Prozentränge für die Rohwerte der Selbsteinschätzung finden sich:

- ◆ für die Gesamtstichprobe in Tabelle A1 des Anhangs (siehe S. 136),
- ◆ für die Substichprobe weiblichen Geschlechts in Tabelle A2 des Anhangs (siehe S. 138) und
- ◆ für die Substichprobe männlichen Geschlechts in Tabelle A3 des Anhangs (siehe S. 140).

4.3.3.2. Normwerte für die Skalen des Fragebogens zur Fremdeinschätzung (SEC-I-OR)

T-Werte und Prozentränge für die Rohwerte der Fremdeinschätzung finden sich:

- ◆ für die Gesamtstichprobe in Tabelle A4 des Anhangs (siehe S. 142),
- ◆ für die Substichprobe weiblichen Geschlechts in Tabelle A5 des Anhangs (siehe S. 144) und
- ◆ für die Substichprobe männlichen Geschlechts in Tabelle A6 des Anhangs (siehe S. 146).

4.3.3.3. Normwerte für den Situational Judgement Test (SEC-SJT)

Für den Situational Judgement Test liegen nur Prozentränge und keine *T*-Werte vor (siehe Abschnitt *Normierung des Situational Judgement Tests*, S. 123). Prozentränge für die Roh-

werte des Situational Judgement Tests finden sich für die Gesamtstichprobe, für die Substichprobe weiblichen Geschlechts und für die Substichprobe männlichen Geschlechts in Tabelle A7 des Anhangs (siehe S. 148).

Literatur

- Ackerman, P. L., Kanfer, R., & Calderwood, C. (2010). Use it or lose it? Wii brain exercise practice and reading for domain knowledge. *Psychology and Aging, 25*(4), 753–766. <https://doi.org/10.1037/a0019277>
- Barchard, K. A., & Russell, J. A. (2006). Bias in consensus scoring, with examples from ability emotional intelligence tests. *Psicothema, 18*(supl.), 49–54. <http://www.psicothema.com/pdf/3275.pdf>
- Brackett, M. A., & Salovey, P. (2006). Measuring emotional intelligence with the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). *Psicothema, 18*(supl.), 34–41. <http://www.psicothema.com/pdf/3273.pdf>
- Bühner, M. (2010). *Einführung in die Test- und Fragebogenkonstruktion* (3. Aufl.). Pearson.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. <https://doi.org/10.1037/h0046016>
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1–22. <https://doi.org/10.1037/h0046743>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Lawrence Erlbaum Associates.
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98–101. <https://doi.org/10.1111/1467-8721.ep10768783>
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement, 6*(4), 475–494. <https://doi.org/10.1177/001316444600600405>
- Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement, 10*(1), 3–31. <https://doi.org/10.1177/001316445001000101>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. <https://doi.org/10.1007/BF02310555>
- Cronbach, L. J. (1990). *Essentials of psychological testing* (5. Aufl.). Harper & Row.
- Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*(3), 391–418. <https://doi.org/10.1177/0013164404266386>
- Dilling, H., & Freyberger, H. J. (Hrsg.). (2012). *Taschenführer zur ICD-10-Klassifikation psychischer Störungen* (6. Aufl.). Huber.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*(3), 215–251. <https://doi.org/10.1037/0033-295X.87.3.215>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.
- Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson

- coefficient twenty is the same for two tests. *Psychometrika*, 34(3), 363–373.
<https://doi.org/10.1007/BF02289364>
- Feldt, L. S., & Kim, S. (2006). Testing the difference between two alpha coefficients with small samples of subjects and raters. *Educational and Psychological Measurement*, 66(4), 589–600. <https://doi.org/10.1177/0013164405282488>
- Flanagan, J. C. (1954). The Critical Incident Technique. *Psychological Bulletin*, 51(4), 327–358. <https://doi.org/10.1037/h0061470>
- Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin*, 138(2), 296–321. <https://doi.org/10.1037/a0026556>
- Häcker, H., Leutner, D., & Amelang, M. (Hrsg.). (1998). Standards für pädagogisches und psychologisches Testen. *Diagnostica, Supplementum 1*.
- Hinsch, R., & Pfingsten, U. (2015). *Gruppentraining sozialer Kompetenzen (GSK)* (6. Aufl.). Beltz.
- Hodzic, S., Scharfen, J., Ripoll, P., Holling, H., & Zenasni, F. (2018). How efficient are emotional intelligence trainings: A meta-analysis. *Emotion Review*, 10(2), 138–148. <https://doi.org/10.1177/1754073917708613>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57(5), 253–270. <https://doi.org/10.1037/h0023816>
- Horn, R. (Hrsg.). (2009). *Raven's Progressive Matrices und Vocabulary Scales: Standard Progressive Matrices (SPM)* (2. Aufl.). Pearson.
- International Test Commission. (2001). *Internationale Richtlinien für die Testanwendung*. https://www.intestcom.org/files/guideline_test_use_german.pdf
- Joseph, D. L., Jin, J., Newman, D. A., & O'Boyle, E. H. (2015). Why does self-reported emotional intelligence predict job performance? A meta-analytic investigation of mixed EI. *Journal of Applied Psychology*, 100(2), 298–342. <https://doi.org/10.1037/a0037681>
- Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: an integrative meta-analysis and cascading model. *The Journal of Applied Psychology*, 95(1), 54–78. <https://doi.org/10.1037/a0017286>
- Kanning, U. P. (2009). *Inventar sozialer Kompetenzen (ISK)*. Hogrefe.
- Kihlstrom, J. F., & Cantor, N. (2011). Social Intelligence. In R. J. Sternberg & S. B. Kaufman (Hrsg.), *The Cambridge Handbook of Intelligence* (S. 564–581). Cambridge University Press. <https://doi.org/10.1017/CBO9780511977244.029>
- Koch, A., & Westhoff, K. (2012). *Task-Analysis-Tools (TAToo) – Schritt für Schritt Unterstützung zur erfolgreichen Anforderungsanalyse*. Pabst.
- Leichte Sprache: Ein Ratgeber*. (2014). Bundesministerium für Arbeit und Soziales. <https://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/a752-ratgeber-leichte-sprache.pdf>

- Libbrecht, N., & Lievens, F. (2012). Validity evidence for the situational judgment test paradigm in emotional intelligence measurement. *International Journal of Psychology*, *47*(6), 438–447. <https://doi.org/10.1080/00207594.2012.682063>
- Lienert, G. A. (1956). Zur Interpretation von Testprofilen. *Diagnostica*, *31*(2).
- Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse*. Beltz PVU.
- Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, *67*(3), 280–296. <https://doi.org/10.1037/0021-9010.67.3.280>
- MacCann, C., Jiang, Y., Brown, L. E. R., Double, K. S., Bucich, M., & Minbashian, A. (2019). Emotional intelligence predicts academic performance: A meta-analysis. *Psychological Bulletin*, *146*(2), 150–186. <https://doi.org/10.1037/bul0000219>
- MacCann, C., & Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion*, *8*(4), 540–551. <https://doi.org/10.1037/a0012746>
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2. Aufl.). Lawrence Erlbaum Associates.
- Martins, A., Ramalho, N., & Morin, E. (2010). A comprehensive meta-analysis of the relationship between emotional intelligence and health. *Personality and Individual Differences*, *49*(6), 554–564. <https://doi.org/10.1016/j.paid.2010.05.029>
- Mattingly, V., & Kraiger, K. (2019). Can emotional intelligence be trained? A meta-analytical investigation. *Human Resource Management Review*, *29*(2), 140–155. <https://doi.org/10.1016/j.hrmr.2018.03.002>
- Mayer, J. D., DiPaolo, M., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment*, *54*(3–4), 772–781. <https://doi.org/10.1080/00223891.1990.9674037>
- Mayer, J. D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence*, *22*(2), 89–113. [https://doi.org/10.1016/S0160-2896\(96\)90011-2](https://doi.org/10.1016/S0160-2896(96)90011-2)
- Mayer, J. D., Salovey, P., Caruso, D. R., & Cherkasskiy, L. (2011). Emotional intelligence. In R. J. Sternberg & S. B. Kaufman (Hrsg.), *The Cambridge Handbook of Intelligence* (S. 528–549). Cambridge University Press. <https://doi.org/10.1017/CBO9780511977244.027>
- McCall, W. A. (1922). *How to measure in education*. MacMillan Co.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. <https://doi.org/10.1016/j.intell.2008.08.004>
- Miao, C., Humphrey, R. H., & Qian, S. (2017). A meta-analysis of emotional intelligence and work attitudes. *Journal of Occupational and Organizational Psychology*, *90*(2), 177–202. <https://doi.org/10.1111/joop.12167>
- O'Boyle, E. H., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. A. (2011). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior*, *32*(5), 788–818. <https://doi.org/10.1002/job.714>

- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396–402. <https://doi.org/10.3758/BF03200807>
- Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*(1), 54–74. <https://doi.org/10.1037/0033-2909.124.1.54>
- Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., Howard, R. J., & Ballard, C. G. (2010). Putting brain training to the test. *Nature*, *465*(7299), 775–778. <https://doi.org/10.1038/nature09042>
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Hrsg.), *The role of constructs in psychological and educational measurement* (Nummer 1958, S. 49–69). Lawrence Erlbaum Associates.
- Payne, J. W. (1994). Thinking aloud: Insights into information processing. *Psychological Science*, *5*(5), 241–248. <https://doi.org/10.1111/j.1467-9280.1994.tb00620.x>
- Petrides, K. V., Siegling, A. B., & Saklofske, D. H. (2016). Theory and measurement of trait emotional intelligence. In U. Kumar (Hrsg.), *The Wiley Handbook of Personality Assessment* (S. 90–103). Wiley. <https://doi.org/10.1002/9781119173489.ch7>
- Rentzsch, K., & Schütz, A. (2009). *Psychologische Diagnostik: Grundlagen und Anwendungsperspektiven*. Kohlhammer.
- Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science*, *26*(6), 515–520. <https://doi.org/10.1177/0963721417712760>
- Sala, G., Tatlidil, K. S., & Gobet, F. (2018). *Video Game Training Does Not Enhance Cognitive Ability : A Comprehensive Meta-Analytic Investigation*. *144*(2), 111–139. <https://doi.org/10.1037/bul0000139>
- Sánchez-Álvarez, N., Berrios Martos, M. P., & Extremera, N. (2020). A meta-analysis of the relationship between emotional intelligence and academic performance in secondary education: A multi-stream comparison. *Frontiers in Psychology*, *11*(1517), 1–11. <https://doi.org/10.3389/fpsyg.2020.01517>
- Sánchez-Álvarez, N., Extremera, N., & Fernández-Berrocal, P. (2016). The relation between emotional intelligence and subjective well-being: A meta-analytic investigation. *Journal of Positive Psychology*, *11*(3), 276–285. <https://doi.org/10.1080/17439760.2015.1058968>
- Schneider, S., Wolf, D., & Schütz, A. (2020). *Workshop zur Diagnostik sozial-emotionaler Kompetenzen: SEC-I und SEC-SJT in der Anwendung*. Otto-Friedrich-Universität Bamberg.
- Schutte, N. S., Malouff, J. M., Thorsteinsson, E. B., Bhullar, N., & Rooke, S. E. (2007). A meta-analytic investigation of the relationship between emotional intelligence and health. *Personality and Individual Differences*, *42*(6), 921–933. <https://doi.org/10.1016/j.paid.2006.09.003>
- Steinmayr, R., Schütz, A., Hertel, J., & Schröder-Abé, M. (2011). *Mayer-Salovey-Caruso Test zur Emotionalen Intelligenz (MSCEIT™)*. Huber.

- Thorndike, E. L. (1920). Intelligence and its uses. *Harper's Magazine*, *140*, 227–235.
- van der Linden, D., Dunkel, C. S., & Petrides, K. V. (2016). The General Factor of Personality (GFP) as social effectiveness: Review of the literature. *Personality and Individual Differences*, *101*, 98–105. <https://doi.org/10.1016/j.paid.2016.05.020>
- van der Linden, D., Pekaar, K. A., Bakker, A. B., Schermer, J. A., Vernon, P. A., Dunkel, C. S., & Petrides, K. V. (2017). Overlap between the general factor of personality and emotional intelligence: A meta-analysis. *Psychological Bulletin*, *143*(1), 36–52. <https://doi.org/10.1037/bul0000078>

Anhang

Tabelle A1

Normwerte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) für Jugendliche und junge Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren (N = 306)

T	SEC-SR	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	T	PR
79	[204]							45]		79	100
78										78	100
77										77	100
76	[198]									76	100
75			25]			25]		44		75	99
74	194]			25]						74	99
73	193									73	99
72								43		72	99
71	192						20]			71	98
70	190 – 191					24		42		70	98
69	189		24							69	97
68	[186 – 188			24						68	97
67	183 – 184]					23				67	96
66	182	20]					19	41		66	95
65	181		23		25]	22				65	94
64	179 – 180			23						64	93
63	178		22				18		20]	63	91
62	175 – 177					21		40		62	89
61	173 – 174	19	21				17			61	87
60	171 – 172			22				39		60	85
59	169 – 170	18			24	20		38		59	83
58	166 – 168									58	80
57	165		20	21		19	16	37		57	77
56	163 – 164	17			23					56	74
55	161 – 162		19					36	19	55	71
54	159 – 160			20		18	15			54	67
53	157 – 158				22			35		53	64
52	155 – 156	16	18	19			14			52	60
51	153 – 154					17		34		51	56
50	151 – 152				21				18	50	52
49	150	15	17	18			13	33		49	48
48	148 – 149				20	16				48	44
47	146 – 147						12	32	17	47	40

(Fortsetzung)

Tabelle A1

<i>T</i>	SEC-SR	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	<i>T</i>	PR
46	145	14	16			15		31		46	36
45	141 – 144			17	19					45	33
44	138 – 140	13				14	11	30	16	44	29
43	135 – 137		15		18			29		43	26
42	133 – 134			16	17		10			42	23
41	132	12				13		28	15	41	20
40	129 – 131		14	15	16					40	17
39	126 – 128							27		39	15
38	124 – 125			14	15	12	9	26	14	38	13
37	122 – 123	11								37	11
36	120 – 121		13		14			25		36	9
35	119			13		11	8	24	13	35	7
34	118				13					34	6
33	[117	10	12			10		23		33	5
32	114 – 115]		11	12	12		7		12	32	4
31	[113	9		11	11	9		22		31	3
30	109]		10							30	3
29	108	8		10	10	8	6	21	11	29	2
28	107		9			[7				28	2
27		[7		[9				20	10	27	1
26	[106						5	[19		26	1
25					[9	[5]				25	1
24	[100]		[8	[7]					[9	24	1
23		[4]					[4	[14]		23	0
22										22	0
21	[59]		[5]	[5]					[7]	21	0

Anmerkung. Sprünge in der Verteilung der Rohwerte der Normstichprobe sind durch eckige Klammern kenntlich gemacht. *T* = *T*-Wert; PR = Prozentrang. SEC-SR = Sozial-emotionale Kompetenz (Selbstbild); EBE-SR = Empathisches Verhalten (Selbstbild); ECO-SR = Kognitive Empathie (Selbstbild); PEA-SR = Wahrnehmung der Emotionen anderer (Selbstbild); PES-SR = Wahrnehmung eigener Emotionen (Selbstbild); REA-SR = Regulation der Emotionen anderer (Selbstbild); RES-SR = Regulation eigener Emotionen (Selbstbild); SOB-SR = Positive Beziehungsgestaltung (Selbstbild); SON-SR = Beachtung sozialer Normen (Selbstbild).

Tabelle A2

Normwerte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) für Jugendliche und junge Erwachsenen weiblichen Geschlechts mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren (N = 111)

T	SEC-SR	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	T	PR
76	[198]		25]	25]			20]	45]		76	100
75										75	99
74						25]				74	99
73										73	99
72	[188]							44		72	99
71			24							71	98
70	[186]							43		70	98
69	183]			24		24	19	42		69	97
68			[23							68	97
67	182									67	96
66	181			23	25]	23	18			66	95
65	[180	20]						41		65	94
64	178]						17			64	93
63	[177		21]	22		22				63	91
62	174 – 175]							40	20]	62	89
61	172 – 173					21	16			61	87
60	[170 – 171				24					60	85
59	[168]	19	20	21				39		59	83
58	165 – 166]					20	15	38		58	80
57		18			23					57	77
56	164		19							56	74
55	163	17		20		19	14	37		55	71
54	[161 – 162				22				19	54	67
53	158 – 159]		18			18	13	36		53	64
52	156 – 157							35		52	60
51	154 – 155	16		19	21					51	56
50	152 – 153					17	12	34		50	52
49	151		17						18	49	48
48	[150	15		18	20	16				48	44
47	147 – 148]						11	33		47	40
46	146								17	46	36
45	145	14	16		19	15	10			45	33
44	143 – 144			17				32		44	29

(Fortsetzung)

Tabelle A2

<i>T</i>	SEC-SR	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	<i>T</i>	PR
43	139 – 142		15		18				16	43	26
42	137 – 138	13			17	14	9	31		42	23
41	134 – 136				16			30		41	20
40	133		14	16				29		40	17
39	[131 – 132				15	13			15	39	15
38	[127 – 128]	12					8			38	13
37	124 – 125]		13	14 – 15	14			28		37	11
36	[122 – 123				13		7	27		36	9
35	119]		12	13	12	12			14	35	7
34	[118	11	11		11	11		[26		34	6
33				12		10	6			33	5
32			10		10					32	4
31	[114]			11		9		24]	[13	31	3
30	108]	[10	9			8	5			30	3
29										29	2
28	[107		[8	[10	[9	[7		[23	[10]	28	2
27										27	1
26							[4			26	1
25										25	1
24	[59]	[4]	[5]	[5]		[5]		[14]	[7]	24	1

Anmerkung. Sprünge in der Verteilung der Rohwerte der Normstichprobe sind durch eckige Klammern kenntlich gemacht. *T* = *T*-Wert; PR = Prozentrang. SEC-SR = Sozial-emotionale Kompetenz (Selbstbild); EBE-SR = Empathisches Verhalten (Selbstbild); ECO-SR = Kognitive Empathie (Selbstbild); PEA-SR = Wahrnehmung der Emotionen anderer (Selbstbild); PES-SR = Wahrnehmung eigener Emotionen (Selbstbild); REA-SR = Regulation der Emotionen anderer (Selbstbild); RES-SR = Regulation eigener Emotionen (Selbstbild); SOB-SR = Positive Beziehungsgestaltung (Selbstbild); SON-SR = Beachtung sozialer Normen (Selbstbild).

Tabelle A3

Normwerte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) für Jugendliche und junge Erwachsenen männlichen Geschlechts mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren (N = 195)

T	SEC-SR	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	T	PR
78	[204]							44]		78	100
77										77	100
76						25]				76	100
75										75	99
74	194]		25]							74	99
73				25]				43		73	99
72	193					24				72	99
71								42		71	98
70	192						20]			70	98
69	191					23				69	97
68	189 – 190		24							68	97
67	[188	20]		24				41		67	96
66	[184]									66	95
65	181 – 182]				25]	22	19			65	94
64	180		23							64	93
63	178 – 179			23				40	20]	63	91
62	[176 – 177	19				21	18			62	89
61	173 – 174]		22					39		61	87
60	171 – 172	18				20		38		60	85
59	170		21	22	24		17			59	83
58	168 – 169					19		37		58	80
57	165 – 167	17	20	21						57	77
56	[164							36	19	56	74
55	160 – 162]				23	18	16			55	71
54	158 – 159		19	20				35		54	67
53	156 – 157	16								53	64
52	154 – 155		18	19	22	17	15	34		52	60
51	152 – 153								18	51	56
50	151	15			21		14	33		50	52
49	149 – 150		17	18		16		32		49	48
48	148				20		13		17	48	44
47	146 – 147	14						31		47	40
46	143 – 145		16	17	19	15		30		46	36

(Fortsetzung)

Tabelle A3

<i>T</i>	SEC-SR	EBE-SR	ECO-SR	PEA-SR	PES-SR	REA-SR	RES-SR	SOB-SR	SON-SR	<i>T</i>	PR
45	139 – 142						12		16	45	33
44	135 – 138	13			18	14		29		44	29
43	133 – 134		15	16			11			43	26
42	132	12				13		28	15	42	23
41	130 – 131			15	17					41	20
40	127 – 129		14					27		40	17
39	125 – 126				16	12	10	26	14	39	15
38	[124	11		14				25		38	13
37	121 – 122]				15					37	11
36	120		13			11	9	24	13	36	9
35	118 – 119	10		13	14			23		35	7
34										34	6
33	[117	9			13	10	8		12	33	5
32	114 – 115]		12	[12				22		32	4
31	[113	8			12					31	3
30			11	10]		9		21		30	3
29	[109]		10			[8			11	29	2
28	107]				11		7	20		28	2
27		[7		[9						27	1
26	[106		[9		10	[5]		[19	10	26	1
25										25	1
24										24	1
23										23	0
22	[100]	[4]		[7]	[9		[6	[14]	[9	22	0

Anmerkung. Sprünge in der Verteilung der Rohwerte der Normstichprobe sind durch eckige Klammern kenntlich gemacht. *T* = *T*-Wert; PR = Prozentrang. SEC-SR = Sozial-emotionale Kompetenz (Selbstbild); EBE-SR = Empathisches Verhalten (Selbstbild); ECO-SR = Kognitive Empathie (Selbstbild); PEA-SR = Wahrnehmung der Emotionen anderer (Selbstbild); PES-SR = Wahrnehmung eigener Emotionen (Selbstbild); REA-SR = Regulation der Emotionen anderer (Selbstbild); RES-SR = Regulation eigener Emotionen (Selbstbild); SOB-SR = Positive Beziehungsgestaltung (Selbstbild); SON-SR = Beachtung sozialer Normen (Selbstbild).

Tabelle A4

Normwerte des Fragebogens zur Fremdeinschätzung (SEC-I-OR) für Jugendliche und junge Erwachsenen mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren (N = 306)

T	SEC-OR	EBE-OR	ECO-OR	PEA-OR	PES-OR	REA-OR	RES-OR	SOB-OR	SON-OR	T	PR
79	[194]						19]	[45]		79	100
78										78	100
77						23]				77	100
76	187]	20]						[43]		76	100
75				25]			18			75	99
74	186									74	99
73	[185		25]			22		41]		73	99
72	183]	19								72	99
71	[182			24	25]	21	17	40		71	98
70	179]									70	98
69	178		24					39		69	97
68	175 – 177	18	23	23						68	97
67	172 – 174					20			20]	67	96
66	[171		22	22	24		16	38		66	95
65	[169]									65	94
64	163 – 165]	17	21	21	23					64	93
63	162					19		37		63	91
62	161									62	89
61	158 – 160				22		15	36	19	61	87
60	157		20	20		18				60	85
59	155 – 156	16			21			35		59	83
58	153 – 154					17	14		18	58	80
57	151 – 152							34		57	77
56	149 – 150		19	19		16		33		56	74
55	147 – 148	15			20				17	55	71
54	144 – 146						13	32		54	67
53	143		18			15				53	64
52	140 – 142	14		18				31	16	52	60
51	137 – 139				19		12			51	56
50	136	13	17			14		30		50	52
49	133 – 135			17				29		49	48
48	131 – 132				18	13	11	28	15	48	44
47	128 – 130	12	16							47	40

(Fortsetzung)

Tabelle A4

<i>T</i>	SEC-OR	EBE-OR	ECO-OR	PEA-OR	PES-OR	REA-OR	RES-OR	SOB-OR	SON-OR	<i>T</i>	PR
46	126 – 127					12		27		46	36
45	124 – 125		15	16	17				14	45	33
44	122 – 123					11	10	26		44	29
43	119 – 121	11	14	15				25		43	26
42	116 – 118				16				13	42	23
41	114 – 115	10				10	9	24		41	20
40	112 – 113		13	14				23	12	40	17
39	110 – 111				15			22		39	15
38	106 – 109	9				9			11	38	13
37	105		12				8	21		37	11
36	[102 – 104			13	14				10	36	9
35	100]	8				8		20		35	7
34	99		11			7		19		34	6
33	96 – 98				13				9	33	5
32	[92 – 95	7		12		6	7	18		32	4
31	89]							17		31	3
30	[87 – 88	6	10	11	12		6			30	3
29	83]							16	[8	29	2
28	[82					[5	5			28	2
27	[79]			10						27	1
26		5	[9					15		26	1
25					[11					25	1
24	[74]			9			[4	14		24	1
23									[6]	23	0
22										22	0
21	[72]	[4		[8				[13		21	0

Anmerkung. Sprünge in der Verteilung der Rohwerte der Normstichprobe sind durch eckige Klammern kenntlich gemacht. *T* = *T*-Wert; PR = Prozentrang. SEC-OR = Sozial-emotionale Kompetenz (Selbstbild); EBE-OR = Empathisches Verhalten (Selbstbild); ECO-OR = Kognitive Empathie (Selbstbild); PEA-OR = Wahrnehmung der Emotionen anderer (Selbstbild); PES-OR = Wahrnehmung eigener Emotionen (Selbstbild); REA-OR = Regulation der Emotionen anderer (Selbstbild); RES-OR = Regulation eigener Emotionen (Selbstbild); SOB-OR = Positive Beziehungsgestaltung (Selbstbild); SON-OR = Beachtung sozialer Normen (Selbstbild).

Tabelle A5

Normwerte des Fragebogens zur Fremdeinschätzung (SEC-I-OR) für Jugendliche und junge Erwachsenen weiblichen Geschlechts mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren (N = 195)

T	SEC-OR	EBE-OR	ECO-OR	PEA-OR	PES-OR	REA-OR	RES-OR	SOB-OR	SON-OR	T	PR
76	[194]			25]		23]	[19]	[43]		76	100
75										75	99
74		20]								74	99
73										73	99
72	[183]		[25]	24		22		41]		72	99
71					25]		17]			71	98
70	[179]									70	98
69		19	23]			21		40		69	97
68	[177]			23						68	97
67			22				16			67	96
66	[171 – 174]				24			39	20]	66	95
65	[169]	18	21	22		20				65	94
64	163]				23			38		64	93
63	162			21			15			63	91
62	161	17			22					62	89
61	159 – 160		20			19		37		61	87
60	157 – 158			20			14		19	60	85
59	156				21	18		36		59	83
58	155	16								58	80
57	[154					17	13	35	18	57	77
56	152]		19	19						56	74
55	150 – 151					16		34		55	71
54	148 – 149				20		12	33	17	54	67
53	[147	15	18							53	64
52	143 – 145]			18		15		32		52	60
51	140 – 142						11		16	51	56
50	137 – 139	14	17		19			31		50	52
49	136					14		30		49	48
48	133 – 135	13		17				29	15	48	44
47	131 – 132		16		18	13	10	28		47	40
46	129 – 130					12			14	46	36
45	125 – 128	12	15	16				27		45	33
44	124						9			44	29

(Fortsetzung)

Tabelle A5

<i>T</i>	SEC-OR	EBE-OR	ECO-OR	PEA-OR	PES-OR	REA-OR	RES-OR	SOB-OR	SON-OR	<i>T</i>	PR
43	[121 – 123				17	11		26	13	43	26
42	[118 – 119]		14							42	23
41	114 – 116]	11		15		10		25	12	41	20
40	113		13		16		8	[24		40	17
39	[112	10						22]	11	39	15
38	107 – 110]			14		9				38	13
37	105 – 106	9			15			21		37	11
36	[103 – 104		12					20	10	36	9
35	[100]				14		7		9	35	7
34	[95 – 98]	8		13		7 – 8		19		34	6
33	[92]		[11				6			33	5
32				12		6			[8	32	4
31	[89]				13		5	[18		31	3
30	[82]									30	3
29				[11				[16]		29	2
28	[79]	[7	[9]		12	[5				28	2
27										27	1
26							[4		[6]	26	1
25										25	1
24	[72]	[4]		[8]	[11			[13]		24	1

Anmerkung. Sprünge in der Verteilung der Rohwerte der Normstichprobe sind durch eckige Klammern kenntlich gemacht. *T* = *T*-Wert; PR = Prozentrang. SEC-OR = Sozial-emotionale Kompetenz (Selbstbild); EBE-OR = Empathisches Verhalten (Selbstbild); ECO-OR = Kognitive Empathie (Selbstbild); PEA-OR = Wahrnehmung der Emotionen anderer (Selbstbild); PES-OR = Wahrnehmung eigener Emotionen (Selbstbild); REA-OR = Regulation der Emotionen anderer (Selbstbild); RES-OR = Regulation eigener Emotionen (Selbstbild); SOB-OR = Positive Beziehungsgestaltung (Selbstbild); SON-OR = Beachtung sozialer Normen (Selbstbild).

Tabelle A6

Normwerte des Fragebogens zur Fremdeinschätzung (SEC-I-OR) für Jugendliche und junge Erwachsenen männlichen Geschlechts mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren (N = 195)

T	SEC-OR	EBE-OR	ECO-OR	PEA-OR	PES-OR	REA-OR	RES-OR	SOB-OR	SON-OR	T	PR
78	187]	20]				23]		[45]		78	100
77										77	100
76							18]			76	100
75										75	99
74	186	19		25]		22		[41]		74	99
73			25]							73	99
72	[185					21	17	39]		72	99
71	[182]			24						71	98
70		18			25]					70	98
69	[178]		24			20			20]	69	97
68	[174 – 175]		23	23						68	97
67	172]							38		67	96
66	[171		22	22	24		16			66	95
65	[169]	17				19				65	94
64	[165]		21	21						64	93
63	163]				23			37	19	63	91
62	[161 – 162							36		62	89
61	158]				22	18				61	87
60	156 – 157	16	20	20			15	35		60	85
59	154 – 155				21	17		34		59	83
58	151 – 153								18	58	80
57	149 – 150					16		33		57	77
56	147 – 148	15	19	19			14			56	74
55	[146				20			32	17	55	71
54	143 – 144]	14				15				54	67
53	141 – 142		18	18			13	31		53	64
52	139 – 140								16	52	60
51	136 – 138	13			19	14		30		51	56
50	[135		17					29		50	52
49	132 – 133]			17		13	12			49	48
48	129 – 131	12			18			28	15	48	44
47	127 – 128		16					27		47	40
46	125 – 126			16	17	12	11			46	36

(Fortsetzung)

Tabelle A6

<i>T</i>	SEC-OR	EBE-OR	ECO-OR	PEA-OR	PES-OR	REA-OR	RES-OR	SOB-OR	SON-OR	<i>T</i>	PR
45	123 – 124	11	15					26	14	45	33
44	121 – 122					11				44	29
43	119 – 120		14	15	16			25		43	26
42	116 – 118	10					10	24	13	42	23
41	114 – 115		13	14		10		23		41	20
40	112 – 113								12	40	17
39	110 – 111	9			15		9	22		39	15
38	106 – 109		12						11	38	13
37	[105			13	14	9		21		37	11
36	[102 – 103]	8				8	8			36	9
35	100]							20	10	35	7
34		7	11		13	7				34	6
33	[99					6		19		33	5
32	[95 – 96]			12				18		32	4
31	89]	6		11	12		7	17	9	31	3
30	88		10							30	3
29	[87									29	2
28	[83]			10		[5	6	16		28	2
27										27	1
26	[79]	[5			[11		5	15		26	1
25										25	1
24			[9						[8	24	1
23										23	0
22	[74]			[9			[4	[14		22	0

Anmerkung. Sprünge in der Verteilung der Rohwerte der Normstichprobe sind durch eckige Klammern kenntlich gemacht. *T* = *T*-Wert; PR = Prozentrang. SEC-OR = Sozial-emotionale Kompetenz (Selbstbild); EBE-OR = Empathisches Verhalten (Selbstbild); ECO-OR = Kognitive Empathie (Selbstbild); PEA-OR = Wahrnehmung der Emotionen anderer (Selbstbild); PES-OR = Wahrnehmung eigener Emotionen (Selbstbild); REA-OR = Regulation der Emotionen anderer (Selbstbild); RES-OR = Regulation eigener Emotionen (Selbstbild); SOB-OR = Positive Beziehungsgestaltung (Selbstbild); SON-OR = Beachtung sozialer Normen (Selbstbild).

Tabelle A7

Normwerte des Situational Judgement Tests (SEC-SJT) für Jugendliche und junge Erwachsenen beider Geschlechter ($N_{\text{gesamt}} = 306$), weiblichen Geschlechts ($n_{\text{weiblich}} = 111$) bzw. männlichen Geschlechts ($n_{\text{männlich}} = 195$) mit (sub-)klinischer kognitiver bzw. psychischer Beeinträchtigung im Alter von 15;0 bis 24;11 Jahren

SEC-SJT	PR _{gesamt}	PR _{weiblich}	PR _{männlich}
10	100	100	100
9	95	94	96
8	80	76	83
7	61	51	67
6	43	32	49
5	30	22	35
4	19	14	23
3	13	9	15
2	8	7	9
1	3	2	4
0	1	0	2

Anmerkung. SEC-SJT = Social-Emotional Competence Situational Judgement Test Rohwert;
PR = Prozentrang.

Tabelle A8

Tabelle A8

Itemschlüssel für die Zuordnung der Items zu den Subskalen des Fragebogens zur Selbsteinschätzung (SEC-I-SR) bzw. Fremdeinschätzung (SEC-I-OR)

Subskala	Item				
Empathisches Verhalten (EBE)	1	10	19	28	
Kognitive Empathie (ECO)	7	16	25	34	41
Wahrnehmung der Emotionen anderer (PEA)	6	15	24	33	40
Wahrnehmung eigener Emotionen (PES)	3	12	21	30	38
Regulation der Emotionen anderer (REA)	2	11	20	29	37
Regulation eigener Emotionen (RES)	4	13	22	31	
Positive Beziehungsgestaltung (SOB)	5	9	14	18	23
	27	32	36	39	
Beachtung sozialer Normen (SON)	8	17	26	35	

Anmerkung. Für die Bezeichnung der Subskalen wird jeweils die Endung -SR für das Selbstbild bzw. -OR für das Fremdbild an die Abkürzung der Subskala angehängt.

Tabelle A9

Werte zur Bestimmung von Vertrauensbereichen für die T-Werte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) bzw. Fremdeinschätzung (SEC-I-OR) in Abhängigkeit von Fragestellung und Vertrauenswahrscheinlichkeit

Skala	α	Einseitige Fragestellung				Zweiseitige Fragestellung			
		68%	90%	95%	99%	68%	90%	95%	99%
SEC-SR	.92	±1	±4	±5	±6	±3	±5	±5	±7
SEC-OR	.96	±1	±3	±3	±5	±2	±3	±4	±5
EBE-SR	.80	±2	±6	±7	±10	±4	±7	±9	±12
ECO-SR	.75	±2	±6	±8	±12	±5	±8	±10	±13
PEA-SR	.74	±2	±7	±8	±12	±5	±8	±10	±13
PES-SR	.77	±2	±6	±8	±11	±5	±8	±9	±12
REA-SR	.82	±2	±5	±7	±10	±4	±7	±8	±11
RES-SR	.75	±2	±6	±8	±12	±5	±8	±10	±13
SOB-SR	.77	±2	±6	±8	±11	±5	±8	±9	±12
SON-SR	.68	±3	±7	±9	±13	±6	±9	±11	±15
EBE-OR	.87	±2	±5	±6	±9	±4	±6	±7	±9
ECO-OR	.87	±2	±5	±6	±8	±4	±6	±7	±9
PEA-OR	.83	±2	±5	±7	±10	±4	±7	±8	±11
PES-OR	.81	±2	±6	±7	±10	±4	±7	±8	±11
REA-OR	.91	±1	±4	±5	±7	±3	±5	±6	±8
RES-OR	.84	±2	±5	±7	±9	±4	±7	±8	±10
SOB-OR	.88	±2	±5	±6	±8	±4	±6	±7	±9
SON-OR	.86	±2	±5	±6	±9	±4	±6	±7	±10

Anmerkung. $N = 306$. -SR bzw. -OR = Selbstbild bzw. Fremdbild; SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen.

Tabelle A10

Tabelle A10

Kritische Differenzen für die T-Werte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) bzw. Fremdeinschätzung (SEC-I-OR) für eine Vertrauenswahrscheinlichkeit von 95% bei zweiseitiger Fragestellung

Skala	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 SEC-SR	8																	
2 SEC-OR	7	5																
3 EBE-SR			12															
4 ECO-SR			13	14														
5 PEA-SR			13	14	14													
6 PES-SR			13	14	14	13												
7 REA-SR			12	13	13	13	12											
8 RES-SR			13	14	14	14	13	14										
9 SOB-SR			13	14	14	13	13	14	13									
10 SON-SR			14	15	15	15	14	15	15	16								
11 EBE-OR			11	12	12	12	11	12	12	13	10							
12 ECO-OR			11	12	12	12	11	12	12	13	10	10						
13 PEA-OR			12	13	13	12	12	13	12	14	11	11	11					
14 PES-OR			12	13	13	13	12	13	13	14	11	11	12	12				
15 REA-OR			11	11	12	11	10	11	11	12	9	9	10	10	8			
16 RES-OR			12	13	13	12	11	13	12	14	11	10	11	12	10	11		
17 SOB-OR			11	12	12	12	11	12	12	13	10	10	11	11	9	10	10	
18 SON-OR			11	12	12	12	11	12	12	13	10	10	11	11	9	11	10	10

Anmerkung. N = 306. -SR bzw. -OR = Selbstbild bzw. Fremdbild; SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen.

Tabelle A11

Kritische Differenzen für die T-Werte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) bzw. Fremdeinschätzung (SEC-I-OR) für eine Vertrauenswahrscheinlichkeit von 95% bei einseitiger Fragestellung

Skala	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 SEC-SR	6																	
2 SEC-OR	6	5																
3 EBE-SR			10															
4 ECO-SR			11	12														
5 PEA-SR			11	12	12													
6 PES-SR			11	11	12	11												
7 REA-SR			10	11	11	11	10											
8 RES-SR			11	12	12	11	11	12										
9 SOB-SR			11	11	12	11	11	11	11									
10 SON-SR			12	12	12	12	12	12	12	13								
11 EBE-OR			10	10	10	10	9	10	10	11	9							
12 ECO-OR			9	10	10	10	9	10	10	11	8	8						
13 PEA-OR			10	11	11	10	10	11	10	11	9	9	10					
14 PES-OR			10	11	11	11	10	11	11	12	9	9	10	10				
15 REA-OR			9	10	10	9	9	10	9	10	8	8	8	9	7			
16 RES-OR			10	11	11	10	10	11	10	11	9	9	9	10	8	9		
17 SOB-OR			9	10	10	10	9	10	10	11	8	8	9	9	8	9	8	
18 SON-OR			10	10	10	10	9	10	10	11	9	9	9	9	8	9	8	9

Anmerkung. N = 306. -SR bzw. -OR = Selbstbild bzw. Fremdbild; SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen.

Tabelle A12

Tabelle A12

Kritische Differenzen für die T-Werte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) bzw. Fremdeinschätzung (SEC-I-OR) für eine Vertrauenswahrscheinlichkeit von 99% bei zweiseitiger Fragestellung

Skala	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 SEC-SR	10																	
2 SEC-OR	9	7																
3 EBE-SR			16															
4 ECO-SR			17	18														
5 PEA-SR			18	18	19													
6 PES-SR			17	18	18	18												
7 REA-SR			16	17	17	17	16											
8 RES-SR			17	18	18	18	17	18										
9 SOB-SR			17	18	18	18	17	18	18									
10 SON-SR			19	19	20	19	18	19	19	21								
11 EBE-OR			15	16	16	16	15	16	16	17	13							
12 ECO-OR			15	16	16	15	14	16	15	17	13	13						
13 PEA-OR			16	17	17	16	15	17	16	18	14	14	15					
14 PES-OR			16	17	17	17	16	17	17	18	15	14	15	16				
15 REA-OR			14	15	15	15	13	15	15	16	12	12	13	13	11			
16 RES-OR			15	16	17	16	15	16	16	18	14	14	15	15	13	15		
17 SOB-OR			15	16	16	15	14	16	15	17	13	13	14	14	12	14	13	
18 SON-OR			15	16	16	16	15	16	16	17	14	13	14	15	12	14	13	14

Anmerkung. N = 306. -SR bzw. -OR = Selbstbild bzw. Fremdbild; SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen.

Tabelle A13

Kritische Differenzen für die T-Werte des Fragebogens zur Selbsteinschätzung (SEC-I-SR) bzw. Fremdeinschätzung (SEC-I-OR) für eine Vertrauenswahrscheinlichkeit von 99% bei einseitiger Fragestellung

Skala	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 SEC-SR	9																	
2 SEC-OR	8	6																
3 EBE-SR			15															
4 ECO-SR			16	16														
5 PEA-SR			16	17	17													
6 PES-SR			15	16	16	16												
7 REA-SR			14	15	15	15	14											
8 RES-SR			16	16	17	16	15	16										
9 SOB-SR			15	16	16	16	15	16	16									
10 SON-SR			17	18	18	17	16	18	17	19								
11 EBE-OR			14	14	15	14	13	14	14	16	12							
12 ECO-OR			13	14	14	14	13	14	14	16	12	12						
13 PEA-OR			14	15	15	15	14	15	15	16	13	13	13					
14 PES-OR			14	15	16	15	14	15	15	16	13	13	14	14				
15 REA-OR			13	14	14	13	12	14	13	15	11	11	12	12	10			
16 RES-OR			14	15	15	15	14	15	15	16	13	12	13	14	12	13		
17 SOB-OR			13	14	14	14	13	14	14	15	12	12	13	13	11	12	12	
18 SON-OR			14	15	15	14	13	15	14	16	12	12	13	13	11	13	12	12

Anmerkung. N = 306. -SR bzw. -OR = Selbstbild bzw. Fremdbild; SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen.

Tabelle A14

Tabelle A14

Mittelwerte, Standardabweichungen, Schiefen, Exzesse und part-whole korrigierte Trennschärfen für die zugehörige Subskala bzw. die Globalskala der Items des Fragebogens zur Selbsteinschätzung (SEC-I-SR)

Skala bzw. Item	M	s	μ_3	μ_4	$r_{I(T-I),sub}$	$r_{I(T-I),global}$
<i>Empathisches Verhalten (EBE-SR)</i>						
01 Ich versuche zu helfen, wenn andere ängstlich sind.	3.64	0.98	-0.31	-0.45	0.63	0.56
10 Ich versuche anderen zu helfen, wenn ich das Gefühl habe, dass sie Hilfe brauchen.	3.98	0.95	-0.62	-0.36	0.61	0.58
19 Ich versuche zu helfen, wenn andere wütend sind.	3.46	1.10	-0.22	-0.83	0.60	0.57
28 Ich versuche zu helfen, wenn andere traurig sind.	4.00	0.93	-0.71	0.14	0.61	0.57
<i>Kognitive Empathie (ECO-SR)</i>						
07 Ich kann verstehen, weshalb andere angespannt sind.	3.43	0.94	-0.31	-0.14	0.43	0.39
16 Ich kann verstehen, weshalb andere fröhlich sind.	3.73	1.04	-0.50	-0.40	0.54	0.55
25 Ich kann verstehen, weshalb andere wütend sind.	3.49	0.92	-0.15	-0.48	0.64	0.53
34 Ich kann verstehen, weshalb andere überrascht sind.	3.39	0.95	-0.28	-0.09	0.53	0.56
41 Ich kann verstehen, weshalb andere sich schämen.	3.35	1.02	-0.30	-0.26	0.46	0.51
<i>Wahrnehmung der Emotionen anderer (PEA-SR)</i>						
06 Ich bemerke, wenn andere schlecht gelaunt sind.	4.01	0.91	-0.95	0.92	0.44	0.38
15 Ich bemerke, wenn andere sich schämen.	3.27	1.08	-0.34	-0.45	0.51	0.45
24 Ich bemerke, wenn andere fröhlich sind.	4.00	0.84	-0.69	0.25	0.49	0.57
33 Ich bemerke, wenn andere ängstlich sind.	3.54	0.99	-0.46	-0.09	0.58	0.56
40 Ich bemerke, wenn andere stolz sind.	3.53	1.05	-0.42	-0.44	0.51	0.54
<i>Wahrnehmung eigener Emotionen (PES-SR)</i>						
03 Ich bemerke, wenn ich stolz bin.	3.69	1.15	-0.46	-0.82	0.39	0.26
12 Ich bemerke, wenn ich traurig bin.	4.26	0.94	-1.27	1.14	0.60	0.42
21 Ich bemerke, wenn ich wütend bin.	4.15	1.09	-1.25	0.75	0.61	0.40
30 Ich bemerke, wenn ich schlecht gelaunt bin.	4.20	1.02	-1.15	0.42	0.61	0.45
38 Ich bemerke, wenn ich überrascht bin.	3.92	1.13	-0.91	-0.01	0.51	0.46
<i>Regulation der Emotionen anderer (REA-SR)</i>						
02 Ich kann andere beruhigen, wenn sie ängstlich sind.	3.40	0.94	-0.23	-0.17	0.58	0.52
11 Ich kann andere beruhigen, wenn sie angespannt sind.	3.24	0.92	-0.07	-0.25	0.63	0.55
20 Ich kann anderen helfen, wenn sie von ihren Gefühlen überwältigt werden.	3.36	1.03	-0.21	-0.47	0.67	0.54
29 Ich kann anderen dabei helfen, mit ihren Gefühlen umzugehen.	3.21	1.07	-0.21	-0.46	0.59	0.52
37 Ich kann andere beruhigen, wenn sie wütend sind.	3.24	1.01	-0.19	-0.33	0.57	0.55

(Fortsetzung)

Regulation eigener Emotionen (RES-SR)

04	Ich kann mich selbst beruhigen, wenn ich angespannt bin.	3.30	1.13	-0.09	-0.84	0.55	0.22
13	Ich kann mich selbst beruhigen, wenn ich ängstlich bin.	3.25	1.13	-0.26	-0.60	0.60	0.25
22	Ich kann mich selbst trösten, wenn ich traurig bin.	3.05	1.23	-0.06	-0.94	0.55	0.22
31	Ich weiß, wie ich mit meinen Gefühlen umgehen kann.	3.59	1.14	-0.35	-0.76	0.48	0.42

Positive Beziehungsgestaltung (SOB-SR)

05	Ich will mich gleich nach einem Streit wieder vertragen.	3.43	1.06	-0.12	-0.87	0.29	0.31
09	Ich unterstütze andere.	3.93	0.86	-0.46	-0.14	0.60	0.60
14	Es beschäftigt mich, wenn ich Streit mit anderen habe.	3.75	1.12	-0.81	0.01	0.24	0.25
18	Ich helfe anderen.	4.08	0.91	-0.79	0.09	0.59	0.59
23	Ich achte darauf, dass sich andere wohlfühlen.	3.79	0.96	-0.73	0.39	0.51	0.54
27	Anderer können sich auf mich verlassen.	4.20	0.87	-1.18	1.54	0.38	0.40
32	Ich versuche herauszufinden, was anderen wichtig ist.	3.42	1.05	-0.42	-0.38	0.59	0.57
36	Ich versuche zu schlichten, wenn andere streiten.	3.37	1.12	-0.25	-0.62	0.47	0.51
39	Ich mache anderen kleine Geschenke.	3.02	1.19	-0.03	-0.83	0.47	0.44

Beachtung sozialer Normen (SON-SR)

08	Ich bedanke mich, wenn andere etwas für mich tun.	4.58	0.69	-1.80	3.55	0.47	0.45
17	Ich achte auf andere.	3.81	1.04	-0.68	-0.11	0.42	0.53
26	Ich entschuldige mich, wenn ich etwas kaputt gemacht habe, das anderen gehört.	4.51	0.85	-1.92	3.44	0.47	0.37
35	Ich entschuldige mich, wenn ich andere verletzt habe.	4.38	0.84	-1.35	1.48	0.55	0.48

Anmerkung. $N = 306$. M = Mittelwert; s = Standardabweichung; μ_3 = Schiefe; μ_4 = Exzess; $r_{I(T-I),sub}$ = part-whole korrigierte Trennschärfe für die zugehörige Subskala; $r_{I(T-I),global}$ = part-whole korrigierte Trennschärfe für die Globalskala.

Tabelle A15

Tabelle A15

Mittelwerte, Standardabweichungen, Schiefen, Exzesse und part-whole korrigierte Trennschärfen für die zugehörige Subskala bzw. die Globalskala der Items des Fragebogens zur Fremdeinschätzung (SEC-I-OR)

Skala bzw. Item	M	s	μ_3	μ_4	$r_{I(T-I),sub}$	$r_{I(T-I),global}$
<i>Empathisches Verhalten (EBE-OR)</i>						
01 Sie/Er versucht zu helfen, wenn andere ängstlich sind.	3.16	0.98	-0.25	-0.60	0.74	0.64
10 Sie/Er versucht anderen zu helfen, wenn sie/er das Gefühl hat, dass sie Hilfe brauchen.	3.58	0.95	-0.46	-0.21	0.71	0.74
19 Sie/Er versucht zu helfen, wenn andere wütend sind.	2.93	0.90	-0.02	-0.55	0.65	0.61
28 Sie/Er versucht zu helfen, wenn andere traurig sind.	3.30	0.90	-0.20	-0.32	0.77	0.70
<i>Kognitive Empathie (ECO-OR)</i>						
07 Sie/Er kann verstehen, weshalb andere angespannt sind.	3.28	0.91	-0.25	-0.47	0.68	0.63
16 Sie/Er kann verstehen, weshalb andere fröhlich sind.	3.68	0.76	-0.52	0.31	0.68	0.63
25 Sie/Er kann verstehen, weshalb andere wütend sind.	3.36	0.84	-0.12	-0.58	0.74	0.65
34 Sie/Er kann verstehen, weshalb andere überrascht sind.	3.25	0.78	0.02	-0.17	0.71	0.61
41 Sie/Er kann verstehen, weshalb andere sich schämen.	3.13	0.88	0.05	-0.47	0.70	0.64
<i>Wahrnehmung der Emotionen anderer (PEA-OR)</i>						
06 Sie/Er bemerkt, wenn andere schlecht gelaunt sind.	3.76	0.76	-0.56	0.46	0.58	0.62
15 Sie/Er bemerkt, wenn andere sich schämen.	3.18	0.77	0.04	-0.16	0.64	0.59
24 Sie/Er bemerkt, wenn andere fröhlich sind.	3.85	0.65	-0.57	0.93	0.65	0.65
33 Sie/Er bemerkt, wenn andere ängstlich sind.	3.24	0.85	-0.16	-0.61	0.67	0.68
40 Sie/Er bemerkt, wenn andere stolz sind.	3.20	0.82	0.04	-0.25	0.65	0.57
<i>Wahrnehmung eigener Emotionen (PES-OR)</i>						
03 Sie/Er bemerkt, wenn sie/er stolz ist.	3.60	0.85	-0.33	-0.18	0.53	0.40
12 Sie/Er bemerkt, wenn sie/er traurig ist.	3.91	0.76	-0.35	-0.14	0.58	0.48
21 Sie/Er bemerkt, wenn sie/er wütend ist.	3.72	0.86	-0.53	0.17	0.62	0.43
30 Sie/Er bemerkt, wenn sie/er schlecht gelaunt ist.	3.82	0.80	-0.40	-0.15	0.67	0.43
38 Sie/Er bemerkt, wenn sie/er überrascht ist.	3.52	0.79	0.01	-0.42	0.63	0.56
<i>Regulation der Emotionen anderer (REA-OR)</i>						
02 Sie/Er kann andere beruhigen, wenn sie ängstlich sind.	2.84	0.94	-0.03	-0.47	0.75	0.69
11 Sie/Er kann andere beruhigen, wenn sie angespannt sind.	2.82	0.91	-0.03	-0.44	0.77	0.69
20 Sie/Er kann anderen helfen, wenn sie von ihren Gefühlen überwältigt werden.	2.68	0.87	0.12	-0.32	0.77	0.70
29 Sie/Er kann anderen dabei helfen, mit ihren Gefühlen umzugehen.	2.56	0.88	0.11	-0.45	0.79	0.74
37 Sie/Er kann andere beruhigen, wenn sie wütend sind.	2.74	0.91	-0.06	-0.59	0.79	0.70

(Fortsetzung)

Regulation eigener Emotionen (RES-OR)

04	Sie/Er kann sich selbst beruhigen, wenn sie/er angespannt ist.	2.95	0.89	-0.18	-0.66	0.70	0.37
13	Sie/Er kann sich selbst beruhigen, wenn sie/er ängstlich ist.	2.93	0.78	0.07	-0.59	0.70	0.36
22	Sie/Er kann sich selbst trösten, wenn sie/er traurig ist.	2.84	0.79	0.00	0.02	0.67	0.44
31	Sie/er weiß, wie sie/er mit ihren/seinen Gefühlen umgehen kann.	2.92	0.93	0.05	-0.57	0.65	0.58
<i>Positive Beziehungsgestaltung (SOB-OR)</i>							
05	Sie/Er will sich gleich nach einem Streit wieder vertragen.	3.38	0.98	-0.12	-0.67	0.51	0.46
09	Sie/Er unterstützt andere.	3.53	0.96	-0.47	-0.12	0.77	0.73
14	Es beschäftigt sie/ihn, wenn sie/er Streit mit anderen hat.	3.96	0.79	-0.47	-0.11	0.38	0.35
18	Sie/Er hilft anderen.	3.60	0.92	-0.32	-0.50	0.78	0.76
23	Sie/Er achtet darauf, dass andere sich wohlfühlen.	3.09	0.90	-0.13	-0.68	0.75	0.75
27	Anderer können sich auf sie/ihn verlassen.	3.56	0.97	-0.26	-0.46	0.64	0.66
32	Sie/Er versucht herauszufinden, was anderen wichtig ist.	2.86	0.91	0.02	-0.41	0.68	0.68
36	Sie/Er versucht zu schlichten, wenn andere streiten.	3.02	1.05	-0.07	-0.62	0.64	0.65
39	Sie/Er macht anderen kleine Geschenke.	2.27	0.93	0.50	-0.11	0.41	0.37
<i>Beachtung sozialer Normen (SON-OR)</i>							
08	Sie/Er bedankt sich, wenn andere etwas für sie/ihn tun.	3.97	0.92	-0.69	0.06	0.77	0.64
17	Sie/Er achtet auf andere.	3.63	0.88	-0.24	-0.60	0.54	0.73
26	Sie/Er entschuldigt sich, wenn sie/er etwas kaputt gemacht hat, das anderen gehört.	3.92	0.98	-0.70	-0.10	0.75	0.62
35	Sie/Er entschuldigt sich, wenn sie/er andere verletzt hat.	3.73	0.99	-0.53	-0.32	0.76	0.63

Anmerkung. $N = 306$. M = Mittelwert; s = Standardabweichung; μ_3 = Schiefe; μ_4 = Exzess.

Tabelle A16

Tabelle A16

Relative Häufigkeiten der gewählten Antwortalternative und part-whole korrigierte Trennschärfen ($r_{I(T-I)}$) der Items des Situational Judgement Tests (SEC-SJT)

Item	Antwortalternative					$r_{I(T-I)}$
	1	2	3	4	5	
1	16.67%	20.59%	61.44%	0.33%	1.63%	.30
2	4.25%	0.98%	10.78%	65.03%	19.93%	.25
3	17.97%	65.36%	0.98%	6.86%	9.48%	.07
4	80.72%	2.61%	3.27%	12.09%	1.31%	.23
5	5.23%	18.63%	5.23%	56.54%	14.38%	.36
6	6.21%	10.46%	67.97%	11.76%	4.58%	.47
7	3.27%	10.13%	76.14%	2.94%	6.86%	.42
8	7.84%	16.01%	59.80%	3.27%	12.09%	.44
9	30.72%	51.96%	5.56%	8.17%	4.25%	.30
10	9.15%	17.65%	66.67%	2.29%	4.25%	.31

Anmerkung. $N = 306$. Antworten nach Konsensus- bzw. Expertenmethode in fett markiert. $r_{I(T-I)}$ = part-whole korrigierte Trennschärfe.

Tabelle A17

Mittelwerte, Standardabweichungen, Schiefen und Exzesse der Globalskala des Situational Judgement Tests (SEC-SJT), des Fragebogens zur Selbsteinschätzung (SEC-I-SR) und des Fragebogens zur Fremdeinschätzung (SEC-I-OR) bzw. der jeweiligen Subskalen

Skala	<i>M</i>	<i>s</i>	μ_3	μ_4
<i>Sozial-emotionale Kompetenz (SEC)</i>				
Situational Judgement Test (SEC-SJT)	6.46	2.31	-0.71	-0.14
Selbsteinschätzung (SEC-SR)	3.68	0.50	-0.32	0.62
Fremdeinschätzung (SEC-OR)	3.30	0.55	-0.17	-0.08
<i>Empathisches Verhalten (EBE)</i>				
Selbsteinschätzung (EBE-SR)	3.77	0.78	-0.41	0.13
Fremdeinschätzung (EBE-OR)	3.24	0.78	-0.31	-0.31
<i>Kognitive Empathie (ECO)</i>				
Selbsteinschätzung (ECO-SR)	3.48	0.69	-0.16	0.15
Fremdeinschätzung (ECO-OR)	3.34	0.68	-0.03	-0.27
<i>Wahrnehmung der Emotionen anderer (PEA)</i>				
Selbsteinschätzung (PEA-SR)	3.67	0.68	-0.51	0.55
Fremdeinschätzung (PEA-OR)	3.44	0.60	-0.06	0.19
<i>Wahrnehmung eigener Emotionen (PES)</i>				
Selbsteinschätzung (PES-SR)	4.04	0.77	-0.82	0.15
Fremdeinschätzung (PES-OR)	3.71	0.62	-0.06	-0.16
<i>Regulation der Emotionen anderer (REA)</i>				
Selbsteinschätzung (REA-SR)	3.29	0.76	-0.23	0.22
Fremdeinschätzung (REA-OR)	2.73	0.78	-0.09	-0.43
<i>Regulation eigener Emotionen (RES)</i>				
Selbsteinschätzung (RES-SR)	3.30	0.87	-0.09	-0.57
Fremdeinschätzung (RES-OR)	2.91	0.70	-0.14	-0.33
<i>Positive Beziehungsgestaltung (SOB)</i>				
Selbsteinschätzung (SOB-SR)	3.67	0.61	-0.43	0.15
Fremdeinschätzung (SOB-OR)	3.25	0.66	-0.24	-0.38
<i>Beachtung sozialer Normen (SON)</i>				
Selbsteinschätzung (SON-SR)	4.32	0.62	-0.99	0.75
Fremdeinschätzung (SON-OR)	3.81	0.79	-0.50	-0.30

Anmerkung. *N* = 306. *M* = Mittelwert; *s* = Standardabweichung; μ_3 = Schiefe; μ_4 = Exzess.

Tabelle A18

(Inter-)Korrelationen der Standard Progressive Matrices, der jeweiligen Globalkala bzw. Subskalen des Situational Judgement Tests (SEC-SJT), des Fragebogens zur Selbsteinschätzung (SEC-I-SR) und des Fragebogens zur Fremdeinschätzung (SEC-I-OR)

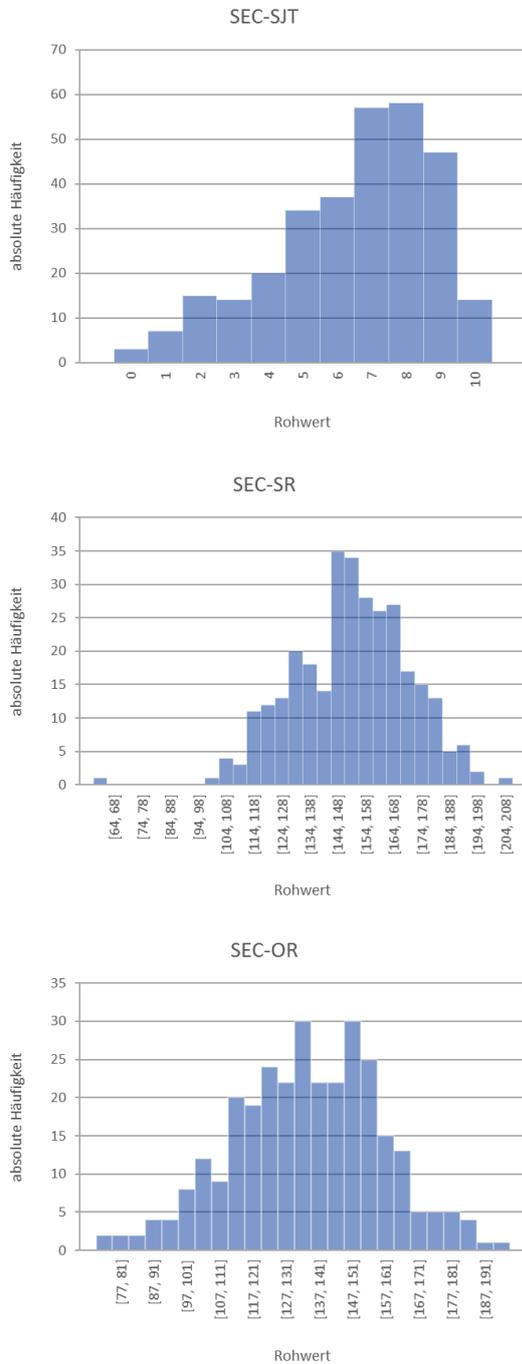
Skala	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 SPM	–																			
2 SEC-SJT ^a	0.33*	(.65)																		
3 SEC-SR ^b	-0.20	0.23***	(.92)																	
4 SEC-OR ^b	0.23	0.27***	0.20***	(.96)																
5 EBE-SR ^b	-0.15	0.22***	0.76***	0.16**	(.80)															
6 ECO-SR ^b	-0.05	0.15*	0.77***	0.14*	0.49***	(.75)														
7 PEA-SR ^b	-0.16	0.17**	0.76***	0.12*	0.48***	0.69***	(.74)													
8 PES-SR ^b	-0.06	0.15**	0.61***	0.18**	0.24***	0.41***	0.43***	(.77)												
9 REA-SR ^b	-0.07	0.14*	0.75***	0.07	0.73***	0.54***	0.53***	0.21***	(.82)											
10 RES-SR ^b	-0.22	-0.07	0.43***	0.08	0.08	0.21***	0.18**	0.43***	0.16**	(.75)										
11 SOB-SR ^b	-0.26	0.27***	0.84***	0.19***	0.75***	0.57***	0.55***	0.30***	0.63***	0.18**	(.77)									
12 SON-SR ^b	0.02	0.26***	0.68***	0.22***	0.49***	0.44***	0.44***	0.43***	0.37***	0.21***	0.59***	(.68)								
13 EBE-OR ^b	0.03	0.18**	0.20***	0.82***	0.23***	0.12*	0.12*	0.10	0.12*	0.02	0.24***	0.17**	(.87)							
14 ECO-OR ^b	0.31	0.24***	0.16**	0.80***	0.13*	0.13*	0.10	0.18**	0.07	0.06	0.10	0.15**	0.51***	(.87)						
15 PEA-OR ^b	0.24	0.26***	0.22***	0.83***	0.16**	0.19***	0.15**	0.21***	0.08	0.04	0.19***	0.23***	0.57***	0.83***	(.83)					
16 PES-OR ^b	0.18	0.18**	0.16**	0.64***	0.12*	0.18**	0.13*	0.16**	0.08	-0.02	0.12*	0.14*	0.40***	0.56***	0.63***	(.81)				
17 REA-OR ^b	0.19	0.22***	0.17**	0.84***	0.15*	0.09	0.10	0.13*	0.10	0.05	0.16**	0.16**	0.77***	0.63***	0.59***	0.41***	(.91)			
18 RES-OR ^b	0.11	0.12*	0.09	0.57***	-0.02	0.04	0.02	0.15*	-0.03	0.29***	0.00	0.05	0.27***	0.46***	0.39***	0.32***	0.51***	(.84)		
19 SOB-OR ^b	0.12	0.26***	0.15**	0.89***	0.13*	0.10	0.07	0.09	0.04	0.04	0.20***	0.19***	0.83***	0.54***	0.63***	0.39***	0.71***	0.38***	(.88)	
20 SON-OR ^b	0.25	0.23***	0.14*	0.81***	0.08	0.09	0.05	0.16**	-0.01	0.06	0.13*	0.25***	0.61***	0.56***	0.61***	0.45***	0.55***	0.39***	0.79***	(.86)

Anmerkung: Reliabilität (Cronbachs Alpha) auf der Diagonalen; SPM = Standard Progressive Matrices; SEC = Sozial-emotionale Kompetenz; EBE = Empathisches Verhalten; ECO = Kognitive Empathie; PEA = Wahrnehmung der Emotionen anderer; PES = Wahrnehmung eigener Emotionen; REA = Regulation der Emotionen anderer; RES = Regulation eigener Emotionen; SOB = Positive Beziehungsgestaltung; SON = Beachtung sozialer Normen; -SJT bzw. -SR bzw. -OR = Leistung bzw. Selbstbild bzw. Fremdbild.

^aN = 36. ^bN = 306.
*p < 0.05. **p < 0.01. ***p < 0.001.

Abbildung A1

Histogramme der Globalskala des Situational Judgement Tests (SEC-SJT), der Globalskala des Fragebogens zur Selbsteinschätzung (SEC-I-SR) und der Globalskala des Fragebogens zur Fremdeinschätzung (SEC-I-OR)

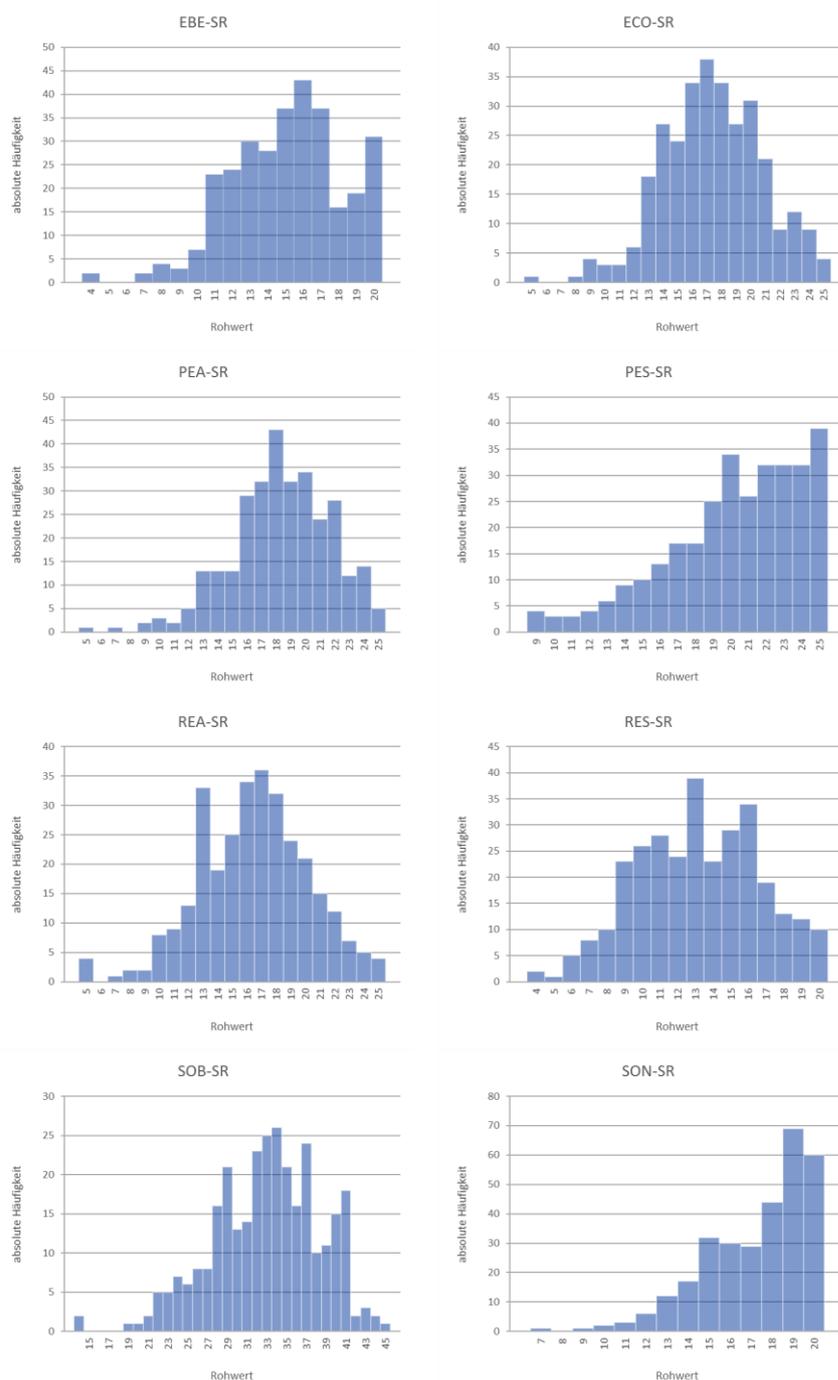


Anmerkung. Häufigkeitsverteilungen der Skalenrohwerter der Globalskala des Situational Judgement Tests, der Globalskala des Fragebogens zur Selbsteinschätzung und der Globalskala des Fragebogens zur Fremdeinschätzung in der Gesamtstichprobe ($N = 306$). SEC-SJT = Sozial-emotionale Kompetenz (Leistung), SEC-SR = Sozial-emotionale Kompetenz (Selbstbild), SEC-OR = Sozial-emotionale Kompetenz (Fremdbild).

Abbildung A2

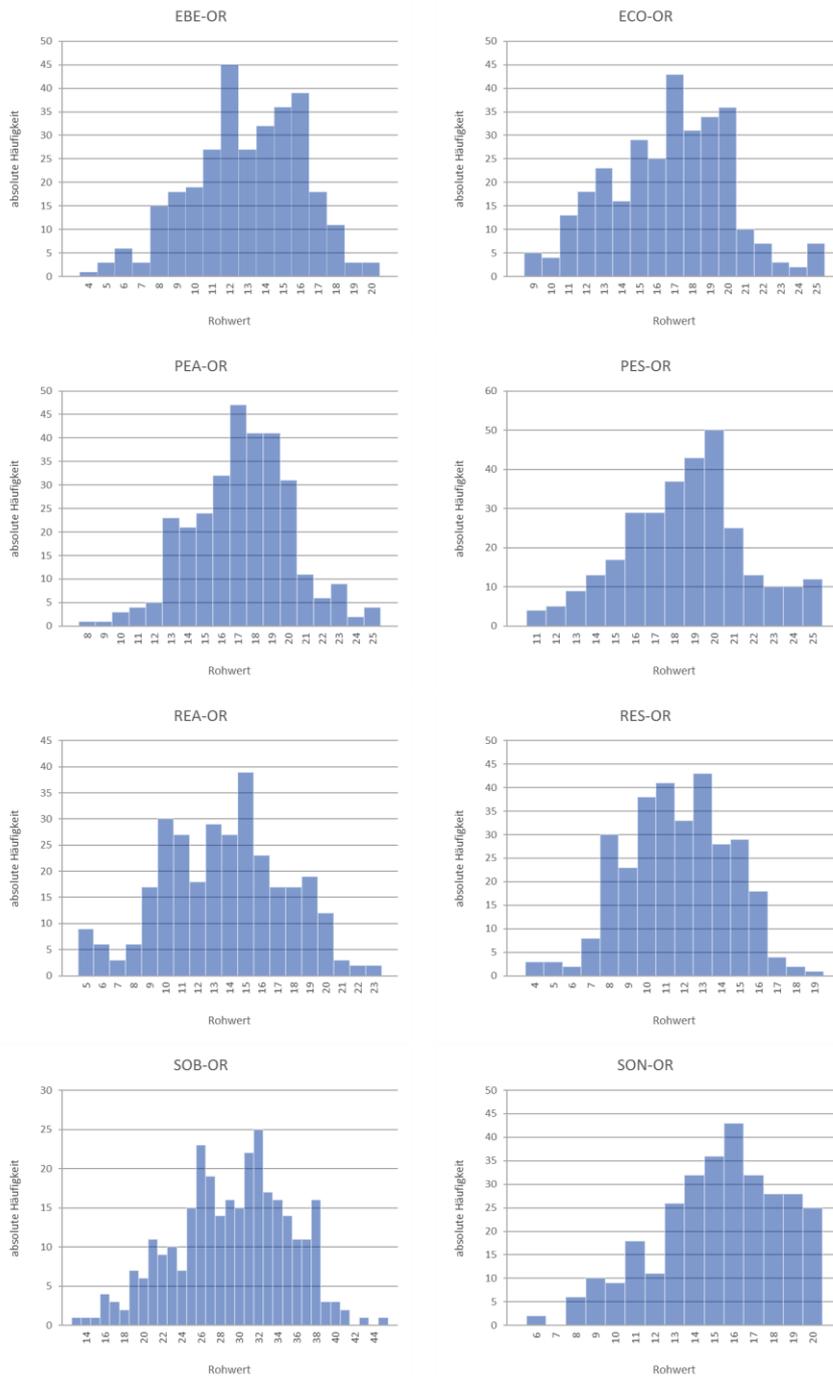
Abbildung A2

Histogramme der Subskalen des Fragebogens zur Selbsteinschätzung (SEC-I-SR)



Anmerkung. Häufigkeitsverteilungen der Subskalenrohre des Fragebogens zur Selbsteinschätzung in der Gesamtstichprobe ($N = 306$). EBE-SR = Empathisches Verhalten (Selbstbild), ECO-SR = Kognitive Empathie (Selbstbild), PEA-SR = Wahrnehmung der Emotionen anderer (Selbstbild), PES-SR = Wahrnehmung eigener Emotionen (Selbstbild), REA-SR = Regulation der Emotionen anderer (Selbstbild), RES-SR = Regulation eigener Emotionen (Selbstbild), SOB-SR = Positive Beziehungsgestaltung (Selbstbild), SON-SR = Beachtung sozialer Normen (Selbstbild).

Abbildung A3

Histogramme der Subskalen des Fragebogens zur Fremdeinschätzung (SEC-I-OR)

Anmerkung. Häufigkeitsverteilungen der Subskalenrohwerter des Fragebogens zur Fremdeinschätzung in der Gesamtstichprobe ($N = 306$). EBE-OR = Empathisches Verhalten (Fremdbild), ECO-OR = Kognitive Empathie (Fremdbild), PEA-OR = Wahrnehmung der Emotionen anderer (Fremdbild), PES-OR = Wahrnehmung eigener Emotionen (Fremdbild), REA-OR = Regulation der Emotionen anderer (Fremdbild), RES-OR = Regulation eigener Emotionen (Fremdbild), SOB-OR = Positive Beziehungsgestaltung (Fremdbild), SON-OR = Beachtung sozialer Normen (Fremdbild).

Das Inventar (SEC-I) und der Situational Judgement Test (SEC-SJT) wurden speziell für die multimethodale Erfassung sozial-emotionaler Kompetenz(en) bei jungen Menschen mit subklinischer oder klinischer kognitiver bzw. psychischer Beeinträchtigung entwickelt. Der Anwendungsbereich liegt vorrangig in der Bestimmung von Förderbedarf. Das Inventar erfasst sozial-emotionale Kompetenz im Selbstbild (Fragebogen zur Selbsteinschätzung) und Fremdbild (Fragebogen zur Fremdeinschätzung). Der Situational Judgement Test erfasst sozial-emotionale Kompetenz als Leistung. Mit dem Inventar werden zusätzlich acht spezifische Subfacetten aus den Bereichen Empathie (Empathisches Verhalten, Kognitive Empathie), Umgang mit Emotionen (Wahrnehmung der Emotionen anderer, Wahrnehmung eigener Emotionen, Regulation der Emotionen anderer, Regulation eigener Emotionen) und prosozialer Orientierung (Positive Beziehungsgestaltung, Beachtung sozialer Normen) erfasst. Die diagnostischen Verfahren können unabhängig voneinander eingesetzt werden. Da Inventar und Situational Judgement Test jedoch aufeinander abgestimmt sind, entfalten die diagnostischen Verfahren ihr Potential insbesondere dann, wenn sie ergänzend zum Einsatz kommen. Mit dem Testmanual liegt eine umfassende Dokumentation von Anwendung, testtheoretischer Fundierung sowie Normierung vor.