Überprüfung der Konstrukt-, Kriteriumsvalidität und Messinvarianz eines Instruments zur Messung von sozialer Unsicherheit

Philipp Nicolay, Simone Weber und Christian Huber

Bergische Universität Wuppertal

Zusammenfassung: Soziale Unsicherheit bei Kindern zeichnet sich durch ein subjektives Unwohlsein in sozialen Situationen und die Sorge vor einer negativen Bewertung durch andere aus. Falls möglich, werden soziale Situationen vermieden oder auf ein Minimum beschränkt. Aufgrund der Unzuverlässigkeit von Fremdurteilen durch Erziehungsberechtigte oder Lehrkräfte haben Selbstbeurteilungsfragebögen in der Diagnostik von sozialer Unsicherheit einen besonderen Stellenwert. Die Social Anxiety Scale for Children (SASC-R-D; Melfsen & Florin, 1997) ist hierbei ein häufig eingesetztes Instrument, deren Validität bisher aber unzureichend untersucht wurde. Die vorliegende Studie untersucht daher anhand von Daten von 997 Kindern die Konstrukt- und Kriteriumsvalidität sowie Messinvarianz über das Geschlecht der Skala. Die Ergebnisse legen nahe, dass es sich um ein valides Instrument handelt, wobei die differenzielle Nützlichkeit der einzelnen Subskalen fragwürdig bleibt. Die Ergebnisse und ihre Implikationen für die praktische Verwendung des Fragebogens werden diskutiert.

Schlüsselwörter: soziale Unsicherheit, SASC-R-D, Validität, Messinvarianz

Assessing the Validity of the Social Anxiety Scale for Children (SASC-R-D)

Abstract: Social unassertiveness in children is characterized by an unease in social situations and the fear of being negatively evaluated by peers. As a result, social interactions are reduced to a bare minimum or avoided altogether. Self-report measures have assumed a special significance for diagnosing social unassertiveness, as teachers and parents often fail to reliably assess internalizing symptoms. In this context, the Social Anxiety Scale for Children (SASC-R-D; Melfsen & Florin, 1997) is an often-used instrument. However, previous findings concerning the underlying dimensional structure are inconclusive and lack proper statistical evaluation. Therefore, this study examines the construct and criterion validity as well as measurement invariance related to sex on the SASC-R-D. A sample of 997 primary-school students participated in this study. Results indicate robust evidence for the overall validity of the SASC-R-D. However, the practical benefit of the proposed subscales seems limited. The results as well as practical implications are discussed.

Keywords: social unassertiveness, social anxiety, SASC-R-D, validity, measurement invariance

Soziale Unsicherheit

Die erfolgreiche Teilnahme an sozialen Interaktionen stellt einen wesentlichen Bestandteil des schulischen Alltags von Kindern dar. Von einigen Schülerinnen und Schülern wird diese Teilnahme an sozialen Situationen als besonders herausfordernd wahrgenommen. Diese Kinder werden im alltäglichen Sprachgebrauch häufig als schüchtern oder zurückhaltend beschrieben. Eine wiederkehrende Bezeichnung für diese Gruppe ist das Konstrukt der sozialen Unsicherheit (Saile & Kison, 2002). Eine Möglichkeit der Früherkennung sozialer Unsicherheit im Kindesalter bietet das Screeningverfahren *Social Anxiety Scale for Children* (SASC-R-D; Melfsen & Florin, 1997). Die

bisher noch ausstehende Überprüfung der Validität und Messinvarianz der SASC-R-D ist das Ziel der vorliegenden Studie.

Theoretische Einordnung des Konstrukts soziale Unsicherheit

Soziale Unsicherheit beschreibt das subjektive Unwohlsein in sozialen Situationen und die Sorge vor einer negativen Bewertung durch andere. Sofern es möglich ist, werden soziale Situationen vermieden oder auf ein Minimum beschränkt (Ahrens-Eipper, Lepow & Nelius, 2010; La Greca & Stone, 1993).

Soziale Unsicherheit wird weder in der ICD-10 (World Health Organization, 2004) noch in dem DSM-V (American Psychiatric Association, 2013) als eigenes Störungsbild aufgeführt. Es sind jedoch einige Parallelen zu Störungsbildern wie der *Störung mit sozialer Ängstlichkeit* oder der *Sozialen Phobie* zu erkennen, weswegen soziale Unsicherheit von einigen Autorinnen und Autoren als subklinische Form ebendieser Störungsbilder erachtet wird (Feiereis, Ahrens-Eipper & Lässig, 2009).

Anders als nach den Diagnosekriterien einer Störung mit sozialer Ängstlichkeit kann soziale Unsicherheit auch nach dem sechsten Lebensjahr erstmalig auftreten, z.B. wenn im schulischen Kontext vermehrt soziale Situationen (Gruppenarbeiten o.ä.) stattfinden. Zudem weisen Kinder mit sozialer Unsicherheit, anders als Kinder mit einer Sozialen Phobie, nicht zwingend einen Leidensdruck oder eine sichtbare Beeinträchtigung in ihrem Alltag auf (Ahrens-Eipper et al., 2010). Dennoch beschreiben Ahrens-Eipper et al. (2010) soziale Unsicherheit im Kindesalter als einen Risikofaktor für die Entwicklung psychischer Störungen (z.B. Angststörungen oder Depression) während der Adoleszenz oder im Erwachsenenalter.

Für soziale Unsicherheit bei Kindern liegen keine Prävalenzraten vor. Dies liegt primär in uneinheitlichen Definitionen und den damit einhergehenden unterschiedlichen Diagnosekriterien begründet (Ahrens-Eipper et al., 2010; Petermann & Petermann, 2015). Ahrens-Eipper et al. (2010) geben einen Überblick über vorhandene Prävalenzraten von verwandten Angststörungen im Sinne der ICD-10 und DSM-V und leiten daraus vorsichtig eine durchschnittliche Prävalenzrate von 7 % für soziale Unsicherheit im Kindesalter ab. Weiter merken sie den Bedarf an aktuellen und gültigen Prävalenzdaten für das Auftreten sozialer Unsicherheit im Kindesalter an.

Kinder mit sozialen Ängsten weisen geringere soziale Kompetenzen auf als Kinder ohne Auffälligkeiten in diesem Bereich (Scharfstein & Beidel, 2015; Stewart & Rubin, 1995).

Des Weiteren zeigten Studien einen negativen Zusammenhang zwischen sozialer Angst und der sozialen Akzeptanz durch Gleichaltrige (La Greca & Stone, 1993; La Greca & Lopez, 1998; Rubin, Chen & Hymel, 1993; Verduin & Kendall, 2008). Da Kinder mit sozialer Unsicherheit soziale Situationen, wenn möglich, vermeiden, ist anzunehmen, dass das erhöhte Risiko unzureichender sozialer Akzeptanz durch Peers auch auf diese Gruppe übertragen werden kann. Interaktionen mit Gleichaltrigen und die damit verbundenen Lernerfahrungen sind jedoch sehr bedeutsam in der kindlichen Entwicklung und stellen einen Schutzfaktor für eine gesunde Entwicklung dar (Inderbitzen-Pisaruk, Clark & Solano, 1992; Laucht, Schmidt & Esser, 2000). Eine Voraussetzung, um betroffene Kinder in Bezug auf ihr Verhalten in sozialen Situationen

unterstützen zu können, ist ein gültiges Diagnoseinstrument.

Diagnostische Herausforderungen bei der Messung sozialer Unsicherheit

Im Zuge der Diagnostik von internalisierenden Störungen und spezifischer von Ängsten im Kindes- und Jugendalter wird grundlegend das Einbeziehen unterschiedlicher Informationsquellen im Sinne eines multimodalen Ansatzes empfohlen. So werden im klinischen Setting neben dem betroffenen Kind selbst auch die Erziehungsberechtigten sowie weitere Bezugspersonen (z.B. Lehrkräfte) als Informationsquelle herangezogen (Essau, 2003; Essau, Conradt & Reiss, 2004; Melfsen & Walitza, 2010). Verschiedene Studien weisen jedoch auf die unzureichenden Übereinstimmungen zwischen diesen Informationsquellen hin (Achenbach, McConaughy & Howell, 1987; McCombs Thomas, Forehand, Armistead, Wierson & Fauber, 1990). Essau et al. (2004) geben zu bedenken, dass insbesondere in einem frühen Stadium von Ängstlichkeit mögliche Symptome von außen (z.B. durch Lehrkräfte) nicht hinreichend wahrgenommen werden können. Da sich Ängstlichkeit in der Regel nicht nur auf den behavioralen oder physiologischen, sondern primär auf den affektiven und kognitiven Ebenen ausdrückt und diese aus einer externen Perspektive nur bedingt wahrnehmbar sind, scheinen Kinder selbst die zuverlässigste Informationsquelle zur Diagnose sozialer Unsicherheit darzustellen (Essau, 2003; Fonseca & Perrin, 2011; Kashdan & Herbert, 2001).

Insbesondere bei jüngeren Kindern gilt es zu beachten, dass die Fähigkeit zur Beantwortung von diagnostischen Fragen von der kognitiven Entwicklung abhängt. Aus diesem Grund wird die Güte von kindlichen Selbstbeurteilungen häufig in Frage gestellt (Essau et al., 2004; Fonseca & Perrin, 2011; Thiels & Schmitz, 2008). Es gibt Hinweise darauf, dass Kinder ab ungefähr sechs Jahren Auskunft über ihre eigene psychische Befindlichkeit geben können (Riley, 2004) und im Laufe der Grundschulzeit eine zunehmend zuverlässigere Selbsteinschätzungsfähigkeit entwickeln (Helmke, 1998; Möller & Trautwein, 2009).

Melfsen und Walitza (2010) weisen darüber hinaus auf die häufig unterschätzte Altersabhängigkeit der Angstsymptomatik hin. Für die Diagnostik sozialer Unsicherheit scheint es daher notwendig, das Vorhandensein entwicklungsspezifischer Ängste, die bei vielen Kindern auftreten und Teil der natürlichen Entwicklung sind, zu berücksichtigen und von Ängsten zu unterscheiden, die nicht oder nicht mehr altersangemessen sind (Muris & Field, 2011; Schneider, 2004; Schneider & Döpfner, 2004).

Während mit zunehmender Schulerfahrung die Angst vor fremden Personen abnimmt, steigt die Angst vor negativer Bewertung durch andere in dieser Zeit an (Schneider, 2004). Dieser natürliche Entwicklungsverlauf sollte bei der Auswertung und Interpretation vorliegender Diagnostikerggebnisse unbedingt berücksichtigt werden.

Vor dem Hintergrund dieser Überlegungen lassen sich folgende Anforderungen an ein Screeningverfahren zur Erfassung von sozialer Unsicherheit ableiten:

1. Aufgrund der unzureichenden Übereinstimmung zwischen Fremd- und Selbsteinschätzung sollte sich das Verfahren an die betroffenen Kinder selbst richten.
2. Um möglichst frühzeitig Kinder mit sozialer Unsicherheit identifizieren zu können, sollte das Verfahren bereits im Primarbereich einsetzbar sein.
3. Das Verfahren sollte zwischen altersangemessenen Ängsten und nicht altersangemessenen (und somit unterstützungsbedürftigen) Ängsten unterscheiden.

Im deutschen Sprachraum werden diese Kriterien durch die deutsche Fassung der SASC-R-D erfüllt. Unterschiedliche Autorinnen und Autoren (Ahrens-Eipper et al., 2010; Melfsen & Florin, 1997; Melfsen, 1998; Petermann & Petermann, 2015) empfehlen die SASC-R-D für die Früherkennung von sozialer Unsicherheit und Angststörungen im Kindesalter.

SASC-R-D

Die SASC-R-D stellt die deutschsprachige Übersetzung des gleichnamigen englischen Instruments dar (La Greca & Stone, 1993) und erfragt ängstliche Kognitionen und Verhaltensweisen in sozialen Situationen im subklinischen Bereich (Melfsen & Walitza, 2010).

Die SASC-R-D ist ein Selbstbeurteilungsfragebogen für Kinder und Jugendliche im Alter von 8 bis 16 Jahren, bei welchem 18 Items in Bezug auf ihre Auftretenshäufigkeit auf einer fünfstufigen Likertskala von *nie* bis *immer* bewertet werden (Melfsen, 1998). Die 18 Items bilden hierbei zwei Skalen zu je neun Items. Die Skala *Angst vor negativer Bewertung* (FNE; Melfsen & Florin, 1997; Melfsen, 1998) beinhaltet Items, die kognitive Aspekte von sozialer Unsicherheit betonen und Gedanken zur Ablehnung durch andere beschreiben (z. B. „Ich glaube, dass andere Jungen und Mädchen sich über mich lustig machen.“). Die Skala *Vermeidung von und Belastung durch soziale Situationen* (SAD; Melfsen & Walitza, 2010) bezeichnet hingegen spezifische soziale Situationen, die vermieden oder als belastend empfunden werden (z. B. „Es fällt mir schwer, andere Jungen und Mädchen zu fragen, ob sie mit mir spielen.“). Diese Skalen decken sich mit der theore-

tischen Definition des Konstrukts soziale Unsicherheit (Melfsen & Florin, 1997).

Die Autorinnen und Autoren geben alters- und geschlechtsspezifische Normwerte für beide Skalen an, die eine getrennte Betrachtung ermöglichen. In der förderpädagogischen Praxis wird hierbei die Auffälligkeit in einer der Skalen als Indikation für eine Intervention verstanden (Ahrens-Eipper et al., 2010; Petermann & Petermann, 2015). Im Forschungskontext werden die Skalen mitunter allerdings auch einzeln eingesetzt (Hartmann et al., 2010) oder zu einem Gesamtscore verrechnet (Keil, Asbrand, Tuschen-Caffier & Schmitz, 2017; Schmidtendorf, Wiedau, Asbrand, Tuschen-Caffier & Heinrichs, 2018).

Entwicklung der SASC-R-D. Die SASC-R wurde bereits mehrfach überarbeitet. Die Überarbeitungen betreffen sowohl die Anzahl der Items und Faktoren als auch deren Zuordnung. Während die erste Fassung der SASC-R (La Greca, Dandes, Wick, Shaw & Stone, 1988) 10 Items umfasste, die den zwei Skalen FNE (6 Items) und SAD (4 Items) zugeordnet wurden, erhöhte sich mit der Veröffentlichung der überarbeiteten Fassung (La Greca & Stone, 1993) die Anzahl der Items auf 18. In diesem Zuge wurde die zweifaktorielle durch eine dreifaktorielle Struktur ersetzt, indem die Skala SAD weiter in eine allgemeine Vermeidung von und Belastung durch soziale Situationen (SAD-G; Beispielitem: „Ich fühle mich unsicher, selbst bei Jungen und Mädchen, die ich sehr gut kenne.“) und eine spezifische Vermeidung von und Belastung durch neue soziale Situationen (SAD-New; Beispielitem: „Ich fühle mich unsicher bei Jungen und Mädchen, die ich nicht kenne.“) ausdifferenzierte wurde. Mit der deutschsprachigen Fassung (Melfsen & Florin, 1997; Melfsen, 1998) wurde die ursprünglich zweifaktorielle Struktur wieder aufgegriffen.

Testgüte der SASC-R-D. Die interne Konsistenz der Subskalen kann als ausreichend bis gut betrachtet werden. Während La Greca und Stone (1993) Cronbachs Alpha Werte für alle drei Skalen $\alpha > .65$ berichten, variieren die Werte in der zweifaktoriellen deutschsprachigen Fassung zwischen $.63 \leq \alpha \leq .79$ (SAD) bzw. $.80 \leq \alpha \leq .87$ (FNE) (Melfsen & Florin, 1997; Melfsen, 1998; Schreier & Heinrichs, 2008). Die Interkorrelationen der Skalen liegen für die dreifaktorielle Struktur im mittleren (.45–.59) Bereich (La Greca & Stone, 1993), variieren in Untersuchungen der zweifaktoriellen Struktur jedoch relativ stark zwischen .34 (Melfsen & Florin, 1997) und .63 (Schreier & Heinrichs, 2008).

Für die deutschsprachige Fassung ermittelte Melfsen (1998) mit einer Teilstichprobe von 40 Schülerinnen und Schülern nach zwei Wochen befriedigende Retest reliabilitäten (FNE = .84; SAD = .74). Allerdings handelte es sich hierbei um eine Stichprobe, die überwiegend ältere Schülerinnen und Schüler ($M = 15.13$, $SD = 0.73$) umfasste.

Erste Ergebnisse zu konvergenten und diskriminanten Validität legen nahe, dass die Skalen der SASC-R-D mittel bis hoch mit anderen Selbstbeurteilungsfragebogen zum Thema Angst korrelieren (Melfsen & Walitza, 2010). Die Übereinstimmung mit dem Fremdurteil durch Lehrkräfte ist erwartungsgemäß gering (Melfsen, 1998). Unklar ist allerdings der Nutzen der Differenzierung in einzelne Subskalen, da sich hier kaum Unterschiede in den Zusammenhängen mit anderen verwendeten Testverfahren zeigen (La Greca & Stone, 1993; Melfsen, 1998).

Zur Untersuchung der Konstruktvalidität der deutschsprachigen Fassung liegen Ergebnisse von Hauptkomponentenanalysen (PCA) mit orthogonaler Rotation nach dem Verfahren Varimax vor (Melfsen & Florin, 1997; Melfsen, 1998; Schreier & Heinrichs, 2008). Dieses Vorgehen ist aus mehreren Gründen als problematisch zu beurteilen. Erstens lässt sich die PCA den hypothesengenerierenden Verfahren der exploratorischen Faktorenanalyse zuordnen (Moosbrugger & Schermelleh-Engel, 2012) und ist dementsprechend kein hypothesenprüfendes Verfahren. In diesem Kontext erklärt sich auch, dass einzelne Items je nach Untersuchung einem anderen Faktor zugeordnet werden (Melfsen & Florin, 1997; Melfsen, 1998; Melfsen & Walitza, 2010). Zweitens sind die Vorannahmen der gewählten Extraktions- (PCA) und Rotationsmethode (orthogonal/Varimax) für die meisten psychologischen Konstrukte nicht dem Gegenstand angemessen und führen zum einen zu einer Überschätzung erklärter Varianz und zum anderen zu einer Unterschätzung der Interkorrelation von Faktoren (Brown, 2015; Moosbrugger & Schermelleh-Engel, 2012). Weiterhin fehlen bisher Untersuchungen der Messinvarianz der SASC-R-D über das Geschlecht, sodass unklar ist, ob die berichteten Mittelwertsunterschiede zwischen Jungen und Mädchen sinnvoll interpretiert werden können. In diesem Sinne kann die Validität der SASC-R-D bis dato als noch nicht gesichert angenommen werden. Ausgehend von dieser Befundlage ist das übergeordnete Ziel dieses Beitrags die Überprüfung der aufgeführten Validitätsaspekte der SASC-R-D.

Ziele und Fragestellung

Zur Überprüfung der Konstruktvalidität soll im Folgenden in einem ersten Schritt die Faktorenstruktur der SASC-R-D mittels konfirmatorischer Faktorenanalyse (CFA) untersucht werden. Aus den dargestellten Veröffentlichungen (La Greca et al., 1988; La Greca & Stone, 1993; Melfsen & Florin, 1997; Melfsen, 1998; Schneider, 2004; Schreier & Heinrichs, 2008), lassen sich zwei mögliche Messmodelle ableiten: ein zweifaktorielles und ein dreifaktorielles. Zu-

sätzlich findet in der Forschungspraxis ein einfaktorielles Modell Anwendung. Im Gegensatz zu den bisherigen Untersuchungen soll die Überprüfung der Konstruktvalidität an einer altershomogenen Stichprobe von Schülerinnen und Schülern der dritten und vierten Jahrgangsstufe erfolgen.

Im Sinne einer Kreuzvalidierung sollen hierfür zuerst die bisher postulierten möglichen Faktorenstrukturen an einem Datensatz evaluiert werden und daran anschließend das am besten passende Modell an einem weiteren Datensatz noch einmal überprüft werden.

In einem zweiten Schritt soll für das zuvor identifizierte Modell die Messinvarianz über das Geschlecht überprüft werden.

Abschließend soll eine Überprüfung der Kriteriumsvalidität vorgenommen werden. In diesem Kontext soll untersucht werden, inwieweit die postulierten Subskalen des zuvor identifizierten Modells einen inkrementellen bzw. differenziellen Mehrwert für die diagnostische Praxis haben.

Methode

Stichprobe

Insgesamt wurden 1 065 Schülerinnen und Schüler aus 16 Schulen in Nordrhein-Westfalen befragt. Nach Ausschluss von 68 Schülern aufgrund fehlender Daten in der SASC-R-D ergab sich eine Stichprobe von $N = 997$ Schülerinnen und Schülern (47 % Mädchen) aus 50 Klassen. In 26 Schulklassen wurde nach Jahrgangsstufen getrennt unterrichtet, wohingegen in 24 Schulklassen 3. und 4. Klassen jahrgangsübergreifend unterrichtet wurden. Insgesamt gehörten 49 % der Schülerinnen und Schüler der dritten Jahrgangsstufe an.

Der Datensatz wurde zufällig in zwei Teildatensätze mit jeweils 25 Schulklassen geteilt. Der erste Teildatensatz bestand aus 474 Schülerinnen und Schülern (49 % Mädchen), wovon 45 % der dritten Jahrgangsstufe angehörten. Der zweite Teildatensatz bestand aus 523 Schülerinnen und Schülern (45 % Mädchen), wovon 53 % der dritten Jahrgangsstufe angehörten.

Durchführung

Die Fragebögen wurden im Rahmen einer größeren Studie eingesetzt und als Paper-Pencil-Befragung durchgeführt. Den Schülerinnen und Schülern wurden dabei die Fragen vorgelesen.

Instrumente

SASC-R-D. Die SASC-R-D (Melfsen & Florin, 1997; Melfsen, 1998) wurde in der oben beschriebenen deutschen Fassung mit 18 Items eingesetzt.

SCAS. Zur Überprüfung der Kriteriumsvalidität wurden zusätzlich die Skalen *Soziale Phobie* und *Generalisierte Angststörung* der deutschen Übersetzung der *Spence Children's Anxiety Scale* (SCAS; Essau, Muris & Ederer, 2002) eingesetzt. Beide Skalen bestehen aus jeweils sechs Items, die auf einer vierstufigen Likert-Skala mit den Endpunkten *niemals* (0) und *immer* (3) beurteilt werden. Im Gegensatz zu anderen Selbstbeurteilungsinstrumenten ist die SCAS spezifisch zur Erfassung von Symptomen bestimmter Angststörungen des DSM-IV entwickelt worden (Hoyer & Margraf, 2003; Melfsen & Walitza, 2010). Die interne Konsistenz im vorliegenden Datensatz liegt für die Skala *Soziale Phobie* bei $\alpha = .72$ und für die Skala *Generalisierte Angststörung* bei $\alpha = .70$.

Datenanalyse

Das Datenmanagement und die Analysen erfolgten mit R 3.6.1 (R Core Team, 2018) unter Anwendung der Pakete *semTools* 0.5–3 (Jorgensen, Pornprasertmanit, Schoemann & Rosseel, 2018), *lavaan* 0.6–6 (Rosseel, 2012) und *BifactorIndicesCalculator* 0.20 (Dueber, 2020).

Für die Abschätzung der Skalenhomogenität wurde, aufgrund der angenommenen mehrfaktoriellen Struktur, das Maß ω gewählt (McDonald, 1999; Schweizer, 2011). Die Prüfung des Messmodells erfolgte über eine CFA mit *Weighted Least Square Means and Variances* Schätzer (WLSMV), um dem kategorialen Antwortformat und den nicht normalverteilten Daten gerecht zu werden (Brown, 2015; Li, 2016). Um die genestete Datenstruktur zu berücksichtigen (Schülerinnen und Schüler in Klassen), wurden darüber hinaus clusterrobuste Standardfehler geschätzt. Die Überprüfung der Modellpassung erfolgte mit den gängigen absoluten (RMSEA, SRMR) und relativen (CFI, TLI) deskriptiven Fit Indizes (Jackson, Gillaspay & Purc-Stephenson, 2009). Als Kriterien für eine gute Modellpassung gelten dabei nach Hu und Bentler (1999) $CFI \text{ \& } TLI \geq .95$, $SRMR \leq .08$ sowie $RMSEA \leq .05$.

Die Überprüfung der geschlechterübergreifenden Messinvarianz des Messmodells wurde ebenfalls mit dem WLSMV-Schätzer im Rahmen einer Mehrgruppen-Faktorenanalyse durchgeführt. Die Prüfung der aufeinanderfolgenden Stufen von Messinvarianz (konfigural, metrisch, skalar) erfolgte mit χ^2 -Differenztests und anhand der von Chen (2007) vorgeschlagenen Heuristik, nach der eine Veränderung von $\Delta CFI > .005$ bzw. $\Delta RMSEA > .01$ für einen substanziell schlechteren Fit spricht.

Alle Modellvergleiche wurde aufgrund des verwendeten WLSMV Schätzers mit dem für die Satorra-Bentler-Korrektur skalierten χ^2 -Differenztest (Satorra, 2000) durchgeführt, der in lavaan analog zur DIFFTEST Funktion von Mplus Verwendung findet (vgl. Svetina, Rutkowski & Rutkowski, 2020).

Zur Überprüfung der inkrementellen Beiträge der Subskalen wurde eine schrittweise Regressionsanalyse durchgeführt.

Ergebnisse

Überprüfung der Konstruktvalidität

In einem ersten Schritt wurden die konkurrierenden Modelle mit CFAs am ersten Teildatensatz überprüft. Wie Tabelle 1 zeigt, passen alle drei Modelle gut zu den Daten ($CFI \text{ \& } TLI \geq .95$, $SRMR \leq .08$, $RMSEA \leq .05$). Das dreifaktorielle Modell weist hierbei insgesamt die besten Fit-Indizes auf ($\chi^2 = 254.50$, $df = 132$, $p < .001$, $CFI = .974$, $TLI = .970$, $SRMR = .046$, $RMSEA = .044$, 90 % CI = [.036; .052]). Der Vergleich der Modelle bestätigt dies: Das dreifaktorielle Modell passt besser als das zweifaktorielle ($\Delta\chi^2 = 24.44$, $\Delta df = 2$, $p \leq .001$) und das einfaktorielle Modell ($\Delta\chi^2 = 71.69$, $\Delta df = 3$, $p \leq .001$) zu den Daten. Im Vergleich zwischen zwei- und einfaktoriellem Modell passt ersteres besser zu den Daten ($\Delta\chi^2 = 36.43$, $\Delta df = 1$, $p \leq .001$).

Die interne Konsistenz der Skalen des dreifaktoriellen Modells liegt im mittleren bis guten Bereich ($\omega_{FNE} = .88$, $\omega_{SAD-G} = .56$, $\omega_{SAD-New} = .63$, $\omega_{total} = .89$). Die standardisierten Faktorladungen sind mit Ausnahme eines Items für alle drei Skalen zufriedenstellend ($.28 \leq \lambda \leq .83$, $M = .62$). Die Interkorrelationen der drei latenten Faktoren liegen im mittleren bis hohen Bereich (.781–.834).

Eine vollständige Darstellung der drei Modelle findet sich im elektronischen Supplement 1.

In einem zweiten Schritt wurde das dreifaktorielle Modell erneut mit dem zweiten Teildatensatz überprüft. Die Passung des Modells ist gut: $\chi^2 = 295.76$, $df = 132$, $p < .001$, $CFI = .973$, $TLI = .968$, $SRMR = .046$, $RMSEA = .049$, 90 % CI = [.041; .056]. Die interne Konsistenz der Skalen ist im mittleren bis guten Bereich ($\omega_{FNE} = .89$, $\omega_{SAD-G} = .63$, $\omega_{SAD-New} = .65$, $\omega_{total} = .90$). Die standardisierten Faktorladungen können für alle drei Skalen als zufriedenstellend beschrieben werden ($.38 \leq \lambda \leq .80$, $M = .65$). Die Interkorrelationen der drei latenten Faktoren liegen im hohen Bereich (.794–.907). Faktorenladungen und Residuen des dreifaktoriellen Messmodells können dem elektronischen Supplement 2 entnommen werden.

Tabelle 1. Modell Fit Indices der drei überprüften Modelle

Model	χ^2	df	CFI	TLI	SRMR	RMSEA	RMSEA 90 % CI
Einfaktorielles Modell	364.33	135	.952	.942	.057	.060	[.053; .067]
Zweifaktorielles Modell	292.21	134	.967	.962	.050	.050	[.042; .058]
Dreifaktorielles Modell	254.50	132	.974	.970	.046	.044	[.036; .052]

Weitergehende Untersuchung der Faktorenstruktur

Da hohe Interkorrelationen zwischen Faktoren auch auf einen übergeordneten Faktor zweiter Ordnung oder ein Bi-Faktormodell hinweisen können (Brown, 2015; Chen, West & Sousa, 2006; Reise, 2012), soll die Passung dieser Modelle im Folgenden explorativ am Gesamtdatensatz untersucht werden.

Im Fall von drei Faktoren erster Ordnung wäre ein Modell zweiter Ordnung genau identifiziert, sodass zwar die Parameter geschätzt, aber kein Modelltest vorgenommen werden könnte. Nach Brown (2015) ist es dennoch sinnvoll, ein solches Modell zu evaluieren, um die Größe der Faktorenladungen höherer Ordnung zu bestimmen. Die Überprüfung eines solchen Modells zeigt, dass die drei Faktoren sehr hoch auf einem übergeordneten Faktor laden ($\gamma_{FNE} = .847$; $\gamma_{SAD-G} = .944$; $\gamma_{SAD-New} = .928$).

Die Überprüfung eines Bi-Faktormodells mit einem globalen Faktor und drei spezifischen Faktoren zeigt, dass dieses gut zu den Daten passt: $\chi^2 = 341.791$, $df = 117$, $p < .001$, CFI = .979, TLI = .973, SRMR = .035, RMSEA = .044, 90 % CI = [.039; .049]. Mit Ausnahme eines Items sind die standardisierten Faktorenladungen auf dem Globalfaktor als zufriedenstellend zu bewerten ($.29 \leq \lambda \leq .70$, $M = .57$). Die Ladungen auf die spezifischen Faktoren fallen für FNE im Vergleich relativ niedrig aus ($.23 \leq \lambda \leq .55$, $M = .37$) und sind für SAD-G ($.04 \leq \lambda \leq .36$, $M = .21$) bzw. SAD-New ($.01 \leq \lambda \leq .38$, $M = .22$) darüber hinaus teilweise nicht signifikant.

Die erklärte gemeinsame Varianz (ECV: explained common variance; Reise, 2012) ist ein Maß, das Auskunft über die Stärke des globalen Faktors geben kann und berechnet sich als Quotient der durch den Globalfaktor erklärten Varianz und der durch den Globalfaktor sowie der spezifischen Faktoren erklärten Varianz. Dieser ECV-Index des Global Faktors liegt mit .77 relativ hoch verglichen mit den ECV-Indizes der spezifischen Faktoren ($.03 \leq ECV \leq .15$). Das hierarchische Omega, als Indikator dafür, inwiefern der Gesamtskalenwert als Maß für einen zugrundeliegenden gemeinsamen Faktor interpretiert werden kann (Reise, 2012), liegt bei $\omega_H = .83$. Im Vergleich dazu befinden sich diese für die Subskalen im niedrigen Bereich ($.08 \leq \omega_H \leq .23$). Dies legt die Interpretation eines Gesamtskalenwerts nah, wohingegen die Subskalenwerte auf wenig gemeinsame Varianz abseits des globalen Faktors zurückzuführen sind. Das Bi-

Faktormodell ist im elektronischen Supplement 3 dargestellt.

Der Vergleich zwischen dem Modell zweiter Ordnung und dem Bi-Faktormodell zeigt, dass das Bifaktor-Modell besser zu den Daten passt ($\Delta\chi^2 = 75.84$, $\Delta df = 15$, $p \leq .001$).

Überprüfung auf Messinvarianz

Bei der Messinvarianzprüfung (siehe Tabelle 2) über das Geschlecht zeigt das konfigurale Invarianzmodell eine gute Passung zu den Daten: $\chi^2 = 551.29$, $df = 264$, $p < .001$, CFI = .971, TLI = .961, SRMR = .047, RMSEA = .047, 90 % CI = [.041; .052], was darauf hindeutet, dass die dreifaktorielle Struktur in beiden Geschlechtern in ähnlicher Weise konzeptualisiert ist. Der Vergleich mit dem metrischen Modell, das zusätzlich Gleichheitsrestriktionen der Faktorenladungen vornimmt, zeigt eine signifikant schlechtere Passung des Modells ($\Delta\chi^2 = 26.82$, $\Delta df = 15$, $p = .03$). Die Veränderung der Fit-Indizes liegt allerdings unter den von Chen (2007) vorgeschlagenen kritischen Werten ($\Delta CFI = .001$, $\Delta RMSEA = .003$). Eine Analyse der Modifikationsindizes legt nah, dass die Faktorenladung des Items „Es fällt mir schwer, andere Jungen und Mädchen zu fragen, ob sie mit mir spielen“ auf dem Faktor SAD-G nicht messinvariant für Jungen und Mädchen ist. Nach dem Aufheben der Gleichheitsrestriktion für die Faktorladung dieses Items zeigt der Vergleich mit dem konfiguralen Messmodell keine Verschlechterung der Passung des Modells ($\Delta\chi^2 = 19.81$, $\Delta df = 14$, $p = .14$), so dass von partieller metrischer Invarianz ausgegangen werden kann. Ein darauffolgender Vergleich mit dem skalaren Invarianzmodell, das Gleichheitsrestriktionen für die Thresholds der Items vornimmt, zeigt eine signifikante Verschlechterung der Passung des Modells ($\Delta\chi^2 = 69.59$, $\Delta df = 51$, $p = .04$), wobei die Veränderung der Fit-Indizes knapp unter den kritischen Werten bleibt ($\Delta CFI = .004$, $\Delta RMSEA = .000$). Vollständige skalare Invarianz für beide Geschlechter kann damit nicht angenommen werden. Die Betrachtung der Modifikationsindizes legt nah, dass die Thresholds des Items „Ich habe Angst, andere zu mir nach Hause einzuladen, weil sie ablehnen könnten“ nicht messinvariant für Jungen und Mädchen sind. Das Aufheben der Gleichheitsrestriktion für die Thresholds des entsprechenden Items zeigt im Vergleich mit dem partiell metrischen Invarianzmodell keine signifikant-

Tabelle 2. Modellgeltungsstatistiken zur Prüfung der Messinvarianz über das Geschlecht

	Modell	χ^2	df	CFI	RMSEA	Modellvergleich	χ^2 diff	Δdf	ΔCFI	$\Delta RMSEA$	p
1	Konfigural	551.29	264	.971	.047						
2	Metrisch	549.16	279	.972	.044	2 vs. 1	26.82	15	.001	.003	.03*
2b	Partiell metrisch	531.34	278	.974	.043	2b vs. 1	19.81	14	.003	.004	.14
3	Skalar	626.33	329	.970	.043	3 vs. 2b	69.59	51	.004	.000	.04*
3b	Partiell skalar	615.21	325	.970	.042	3b vs. 2b	47.75	47	.004	.001	.44

Anmerkung: * p < .05

Tabelle 3. Pearson-Korrelationen zwischen den Skalen-Werten der ein-, zwei- und dreifaktoriellen Variante (1F, 2F, 3F) der SASC-R-D sowie den Skalen Generalisierte Angst und Soziale Phobie der deutschen Version der Spence Children's Anxiety Scale (SCAS-D; Essau, et al. 2000)

Skala	1	2	3	4	5	6	7	8	M	SD
Social Anxiety Scale for Children (SASC-R-D)										
1 1F	1								21.56	12.93
2 2F FNE	.93**	1							10.67	7.78
3 2F SAD	.89**	.65**	1						10.89	6.43
4 3F FNE	.91**	.99**	.64**	1					9.52	7.30
5 3F SAD-G	.76**	.58**	.83**	.57**	1				3.62	3.06
6 3F SAD-New	.83**	.63**	.90**	.59**	.53**	1			8.42	4.76
Spence Children's Anxiety Scale (SCAS-D)										
7. Generalisierte Angst	.57**	.55**	.50**	.54**	.41**	.46**	1		5.69	3.22
8. Soziale Phobie	.68**	.66**	.57**	.66**	.48**	.52**	.59**	1	5.68	3.77

Anmerkungen: Signifikante Korrelationen gekennzeichnet mit ** (p < .001); FNE = Skala Angst vor negativer Evaluation; SAD = Skala Vermeidung von und Belastung durch soziale Situationen; SAD-G = Skala Allgemeine Vermeidung von und Belastung durch soziale Situationen; SAD-New = Skala Vermeidung von und Belastung durch neue Situationen.

te Verschlechterung der Passung des Modells ($\Delta\chi^2 = 58.14$, $\Delta df = 47$, $p = .13$). Die Veränderung der Fit-Indizes ($\Delta CFI = .004$, $\Delta RMSEA = .001$) liegt auch unterhalb der kritischen Werte, sodass partielle skalare Invarianz über die Geschlechter angenommen werden kann.

Vor diesem Hintergrund der partiellen skalaren Invarianz lassen sich Mittelwertsunterschiede auf den manifesten Skalenwerten interpretieren (siehe elektronisches Supplement 4). Mädchen weisen auf allen drei Subskalen signifikant höhere Werte auf (Mann-Whitney-U-Test; $p < .001$). Die Effektstärke dieser Unterschiede ist als niedrig einzuschätzen ($.158 \leq r \leq .297$).

Kriteriumsvalidität

Zur Überprüfung der Kriteriumsvalidität wurden die Zusammenhänge der Skalenwerte der ein-, zwei- und dreifaktoriellen Variante der SASC-R-D mit den Skalen *Generalisierte Angststörung* und *soziale Phobie* der SCAS (Essau et al., 2002) im Gesamtdatensatz berechnet (siehe Tabelle 3).

Alle Skalen korrelieren signifikant und in mittlerer Höhe sowohl mit der Skala *Soziale Phobie* ($.48 \leq r \leq .68$) als auch der Skala *Generalisierte Angst* ($.41 \leq r \leq .57$) der SCAS.

Den höchsten Zusammenhang mit beiden Skalen zeigt dabei der Gesamtskalenwert der einfaktoriellen Lösung, wohingegen die Skala SAD-G der dreifaktoriellen Lösung mit beiden Subskalen der SCAS jeweils die niedrigsten Zusammenhänge aufweist. Die Skala FNE zeigt sowohl in der zwei- als auch in der dreifaktoriellen Lösung jeweils höhere Zusammenhänge mit den zwei Außenkriterien als SAD bzw. SAD-G und SAD-New. Für diese zwei Subskalen der dreifaktoriellen Lösung zeigen sich nur geringe Unterschiede im Zusammenhang mit den jeweiligen Skalen der SCAS-D.

Zur Überprüfung des inkrementellen Vorhersagewerts der Subskalen des dreifaktoriellen Modells wurde eine schrittweise Regressionsanalyse auf die Skala *Soziale Phobie* der SCAS-D berechnet, bei der die latenten Faktorenscores drei Subskalen in absteigender Reihenfolge (FNE, SAD-New, SAD-G) der berechneten Korrelationen mit der Skala Soziale Phobie in das Modell aufgenommen wurden. Diese zeigt, dass lediglich die Hinzunahme der Subskala SAD-G das Modell signifikant verbessert [$F_{\text{change}}(2,978) = 22.76$, $p < .001$], wobei der Zuwachs an erklärter Varianz (korrigiertes R^2) eher gering ausfällt ($R^2_1 = .439$; $R^2_2 = .451$). Eine analoge Regressionsanalyse auf die Skala

Generalisierte Angst zeigt ebenfalls nur eine signifikante Verbesserung des Modells durch die Hinzunahme der Subskala SAD-New [$F_{\text{change}}(2,970) = 25.76, p < .001$] bei geringem Zuwachs der erklärten Varianz ($R^2_1 = .306$; $R^2_2 = .324$).

Diskussion

Das Ziel des vorliegenden Beitrags lag in der Überprüfung der Validität der SASC-R-D an einer großen, altershomogenen Stichprobe. Obwohl die SASC-R-D in Forschung und Praxis eingesetzt wird, fehlt es bis dato an einer vollständigen testtheoretischen Absicherung. Da sich das subklinische Phänomen der sozialen Unsicherheit bei fehlender Identifikation und somit auch ausbleibender Förderung jedoch zu einer Angststörung manifestieren könnte und Angststörungen bei Kindern und Jugendlichen zu massiven Einschränkungen des alltäglichen Lebens führen können (Melfsen & Walitza, 2010), ist eben diese testtheoretische Absicherung der SASC-R-D obligatorisch.

Die Überprüfung der Konstruktvalidität der SASC-R-D zeigt, dass diese für die vorliegende Stichprobe von Schülerinnen und Schüler der dritten und vierten Jahrgangsstufe angenommen werden kann. Eine inferenzstatistische Prüfung legt nahe, dass ein dreifaktorielles Modell besser zu den Daten passt, auch wenn die Unterschiede der Model Fit Indices eher gering ausfallen. Das dreifaktorielle Modell konnte auch an einem zweiten Datensatz kreuzvalidiert werden.

Auf Ebene der Items zeigte sich für die geprüften Modelle im ersten Datensatz, dass jeweils die Entfernung eines spezifischen Items mit niedriger Faktorladung („Ich spreche nur mit Jungen und Mädchen, die ich gut kenne“) den Model-Fit verbessern könnte. Hiervon wurde an dieser Stelle jedoch abgesehen, da das übergeordnete Ziel dieser Studie nicht in einer Überarbeitung des Fragebogens bestand. Zudem kann die Faktorladung dieses Items in der zur Kreuzvalidierung verwendeten zweiten Stichprobe als akzeptabel betrachtet werden.

Aufgrund der hohen Interkorrelationen zwischen den Faktoren wurden weiterführende Faktorenanalysen durchgeführt, die einen Faktor zweiter Ordnung bzw. ein Bi-Faktormodell zugrunde legen. Dabei scheint das Bi-Faktormodell am besten geeignet, um die Struktur der SASC-R-D zu erklären. Die ermittelten hierarchischen Omega-Werte legen nahe, dass die Bildung eines Gesamtscores über alle Skalen hinweg passender als die getrennte Betrachtung der einzelnen Subskalen ist (Brown, 2015). Passend hierzu zeigen sich keine substanziellen Unterschiede im Zusammenhang der einzelnen Subskalenwer-

te mit den hier verwendeten externen Maßen für eine *Soziale Phobie* und eine *Generalisierte Angststörung*. Weitere Analysen zeigten darüber hinaus, dass der inkrementelle Vorhersagewert speziell für die zwei SAD-Subskalen eher gering ausfällt. Die Bildung eines Gesamtscores erscheint darüber hinaus auch vor dem Hintergrund plausibel, dass die praktische Anwendung der SASC-R-D als Screening im schulischen Kontext (Ahrens-Eipper et al., 2010; Petermann & Petermann, 2015) wahrscheinlich keine differenzierte Förderung in Abhängigkeit von den jeweiligen Skalenergebnissen nach sich ziehen würde.

Verbunden mit dem Einsatz der SASC-R-D in der schulischen Praxis ist ebenfalls die Frage nach der Verwendung von geschlechtsspezifischen Normwerten. Die Befundlage ist diesbezüglich uneinheitlich. Während Saile und Kison (2002) von ausbleibenden Unterschieden in der Selbstbeurteilung von sozialer Unsicherheit berichten, konnten Melfsen und Florin (1997) signifikante Unterschiede in der Selbstbeurteilung zwischen Jungen und Mädchen nachweisen. Mädchen aus der vorliegenden Stichprobe erreichten ebenfalls signifikant höhere Werte auf den drei Subskalen der SASC-R-D als Jungen. In diesem Zusammenhang konnte eine Überprüfung der Messäquivalenz zeigen, dass das dreifaktorielle Modell der SASC-R-D ein weitestgehend invariantes Maß für Jungen und Mädchen ist (partielle skalare Invarianz).

Die Verwendung geschlechtsspezifischer Normwerte, wie sie von Melfsen (1998) vorgelegt und von Petermann und Petermann (2015) zur Anwendung vorgeschlagen wurden, führt allerdings dazu, dass bei gleich beschriebener Belastung Jungen mitunter als auffällig, Mädchen jedoch als unauffällig wahrgenommen werden. Vor dem Hintergrund eines präventiven Ansatzes und der höheren Prävalenz von Angststörungen bei Frauen (McLean, Asnaani, Litz & Hofmann, 2011) erscheint dieses Vorgehen problematisch. Daher plädieren wir für die Verwendung einer kriterialen, geschlechtsunabhängigen Norm. Diese kriteriale Norm sollte in einem weiteren Schritt erarbeitet werden.

Die vorliegende Studie weist einige Limitationen auf. Die teilnehmenden Schülerinnen und Schüler stammten zwar aus insgesamt 50 verschiedenen Klassen, jedoch befanden sich alle teilnehmenden Schulen in Nordrhein-Westfalen. Eine Generalisierung der Ergebnisse auf den gesamten deutschsprachigen Raum ist somit aufgrund möglicher sprachlicher sowie demographischer Unterschiede nicht uneingeschränkt möglich. Zu bedenken ist weiter, dass für die Studie ausschließlich Daten aus Perspektive der Schülerinnen und Schüler selbst erhoben wurden. Bezugspersonen, wie z. B. die Lehrkräfte, wurden nicht befragt. Für die vorliegende Stichprobe liegen über das Geschlecht und die Klassenzugehörigkeit hinaus keine weiteren Informationen vor. Um das Konstrukt der so-

zialen Unsicherheit genauer zu spezifizieren, sollten in weiteren Studien mögliche Einflussvariablen und/oder Zusammenhänge mit anderen Merkmalen untersucht werden.

Darüber hinaus sollte, insbesondere vor dem Hintergrund des Alters der Zielgruppe und der hierfür fehlenden Befunde, die Retestreliabilität der SASC-R-D anhand einer altershomogenen Stichprobe aus dem Primarbereich erneut in den Blick genommen werden. Sofern die SASC-R-D sich auch im Bereich der Retestreliabilität als gut erweist, könnte man dieses Instrument nutzen, um die derzeit noch bestehende Forschungslücke nach Prävalenzraten für soziale Unsicherheit zumindest im Primarbereich zu schließen.

Im Sinne eines präventiven Ansatzes ist es notwendig, Kinder mit Risikofaktoren so früh wie möglich zu erkennen, um Maßnahmen zur Förderung anbieten zu können. Der vorliegende Beitrag zeigt, dass die SASC-R-D als Screening zur frühen Identifikation von sozialer Unsicherheit im schulischen Kontext genutzt werden kann. Das eingangs beschriebene und häufig empfohlene multimodale Vorgehen ist wünschenswert, erfordert jedoch ein erhöhtes Maß an zeitlichen und personellen Ressourcen. Da soziale Unsicherheit womöglich primär im schulischen Kontext auftritt, erscheint der Einbezug der Erziehungsberechtigten nicht zwingend zuverlässig und notwendig. Es ist anzunehmen, dass die Klassenlehrkraft, insbesondere im Primarbereich, einen guten Einblick in den schulischen Alltag ihrer Schülerinnen und Schüler hat. Um der Forderung eines multimodalen Vorgehens gerecht zu werden, könnte es daher sinnvoll sein, die Lehrkraft in den weiteren diagnostischen Prozess von Schülerinnen und Schülern, die im Screening auffällig waren, miteinzubeziehen.

Elektronische Supplemente (ESM)

Die elektronischen Supplemente sind mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1026/0012-1924/a000270>

ESM 1. Standardisierte Faktorenladungen und Residualvarianzen der am ersten Teildatensatz überprüften Messmodelle.

ESM 2. Standardisierte Faktorenladungen und Residualvarianzen des dreifaktoriellen Messmodells.

ESM 3. Standardisierte Faktorenladungen des Bi-Faktormodells.

ESM 4. Mittelwertvergleich zwischen Jungen und Mädchen für die manifesten Skalenwerte der drei Subskalen.

Literatur

- Achenbach, T. M., McConaughy, S. H. & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213–232. <https://doi.org/10.1037/0033-2909.101.2.213>
- Ahrens-Eipper, S., Lepow, B. & Nelius, K. (2010). *Mutig werden mit Til Tiger. Ein Trainingsprogramm für sozial unsichere Kinder* (Therapeutische Praxis, 2., erw. Aufl.). Göttingen: Hogrefe.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Methodology in the social sciences, 2nd ed). New York: The Guilford Press.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 464–504. <https://doi.org/10.1080/10705510701301834>
- Chen, F. F., West, S. G. & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate behavioral research*, 41, 189–225. https://doi.org/10.1207/s15327906mbr4102_5
- Dueber, D. M. (2020). *BifactorIndicesCalculator: Bifactor Indices Calculator* [Computer software]. Verfügbar unter: <https://cran.r-project.org/package=BifactorIndicesCalculator>
- Essau, C. A. (2003). *Angst bei Kindern und Jugendlichen* (Bd. 2398). München: Reinhardt. Verfügbar unter: <http://www.socialnet.de/rezensionen/isbn.php?isbn=978-3-8252-2398-4>
- Essau, C. A., Conradt, J. & Reiss, B. (2004). Klassifikation, Epidemiologie und diagnostisches Vorgehen. In S. Schneider (Hrsg.), *Angststörungen bei Kindern und Jugendlichen: Grundlagen und Behandlung* (S. 79–102). Berlin: Springer.
- Essau, C. A., Muris, P. & Ederer, E. M. (2002). Reliability and validity of the Spence Children's Anxiety Scale and the Screen for Child Anxiety Related Emotional Disorders in German children. *Journal of Behavior Therapy and Experimental Psychiatry*, 33 (1), 1–18. [https://doi.org/10.1016/S0005-7916\(02\)00005-8](https://doi.org/10.1016/S0005-7916(02)00005-8)
- Feiereis, E. U., Ahrens-Eipper, S. & Lässig, W. (2009). Auswirkungen eines Klinikaufenthaltes auf die soziale Unsicherheit im Kindesalter. *Kindheit und Entwicklung*, 18, 13–20. <https://doi.org/10.1026/0942-5403.18.1.13>
- Fonseca, A. C. & Perrin, S. (2011). The clinical phenomenology and classification of child and adolescent anxiety. In A. P. Field & W. K. Silverman (Eds.), *Anxiety Disorders in Children and Adolescents* (2nd ed., pp. 25–55). Cambridge: Cambridge University Press.
- Hartmann, T., Zahner, L., Pühse, U., Schneider, S., Puder, J. J. & Kriemler, S. (2010). Physical activity, bodyweight, health and fear of negative evaluation in primary school children. *Scandinavian journal of medicine & science in sports*, 20, 27–34. <https://doi.org/10.1111/j.1600-0838.2009.00888.x>
- Helmke, A. (1998). Vom Optimisten zum Realisten? Zur Entwicklung des Fähigkeitsselbstkonzeptes vom Kindergarten bis zur 6. Klassenstufe. In F. E. Weinert (Hrsg.), *Entwicklung im Kindesalter* (S. 115–132). Weinheim: Beltz Psychologie Verlags Union.
- Hoyer, J. & Margraf, J. (2003). Fragebögen und Ratingskalen für die Angstdiagnostik bei Kindern. In J. Hoyer & J. Margraf (Hrsg.), *Angstdiagnostik* (S. 517–586). Berlin, Heidelberg: Springer.
- Hu, L.-t. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55. <https://doi.org/10.1080/10705519909540118>

- Inderbitzen-Pisaruk, H., Clark, M. L. & Solano, C. H. (1992). Correlates of loneliness in midadolescence. *Journal of Youth and Adolescence*, 21, 151–167. <https://doi.org/10.1007/BF01537334>
- Jackson, D. L., Gillaspay, J. A. & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis. An overview and some recommendations. *Psychological Methods*, 14, 6–23. <https://doi.org/10.1037/a0014694>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M. & Rosseel, Y. (2018). *semTools: Useful tools for structural equation modeling* [Computer software]. Verfügbar unter: <https://CRAN.R-project.org/package=semTools>
- Kashdan, T. B. & Herbert, J. D. (2001). Social anxiety disorder in childhood and adolescence: Current status and future directions. *Clinical Child and Family Psychology Review*, 4, 37–61. <https://doi.org/10.1023/A:1009576610507>
- Keil, V., Asbrand, J., Tuschen-Caffier, B. & Schmitz, J. (2017). Children with social anxiety and other anxiety disorders show similar deficits in habitual emotional regulation. Evidence for a transdiagnostic phenomenon. *European Child & Adolescent Psychiatry*, 26, 749–757. <https://doi.org/10.1007/s00787-017-0942-x>
- La Greca, A. M., Dandes, S. K., Wick, P., Shaw, K. & Stone, W. L. (1988). Development of the Social Anxiety Scale for Children. Reliability and concurrent validity. *Journal of Clinical Child Psychology*, 17, 84–91. https://doi.org/10.1207/s15374424jccp1701_11
- La Greca, A. M. & Lopez, N. (1998). Social anxiety among adolescents: Linkages with peer relations and friendships. *Journal of Abnormal Child Psychology*, 26, 83–94.
- La Greca, A. M. & Stone, W. (1993). Social Anxiety Scale for Children-Revised: Factor structure and concurrent validity. *Journal of Clinical Child Psychology*, 22, 17–27.
- Laucht, M., Schmidt, M. H. & Esser, G. (2000). Risiko- und Schutzfaktoren in der Entwicklung von Kindern und Jugendlichen. *Frühförderung interdisziplinär*, 19, 97–108.
- Li, C.-H. (2016). Confirmatory factor analysis with ordinal data. Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48, 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- McCombs Thomas, A., Forehand, R., Armistead, L., Wierson, M. & Fauber, R. (1990). Cross-informant consistency in externalizing and internalizing problems in early adolescence. *Journal of Psychopathology and Behavioral Assessment*, 12, 255–262. <https://doi.org/10.1007/BF00960622>
- McDonald, R. P. (1999). *Test Theory: A Unified Treatment* (1st ed.). London: Taylor and Francis.
- McLean, C. P., Asnaani, A., Litz, B. T. & Hofmann, S. G. (2011). Gender differences in anxiety disorders. Prevalence, course of illness, comorbidity and burden of illness. *Journal of Psychiatric Research*, 45, 1027–1035. <https://doi.org/10.1016/j.jpsychires.2011.03.006>
- Melfsen, S. (1998). Die deutsche Fassung der Social Anxiety Scale for Children Revised (SASC-R-D): Psychometrische Eigenschaften und Normierung. *Diagnostica*, 44, 153–163.
- Melfsen, S. & Florin, I. (1997). Ein Fragebogen zur Erfassung sozialer Angst bei Kindern (SASC-R-D). *Kindheit und Entwicklung*, 6, 224–229.
- Melfsen, S. & Walitz, S. (2010). Skalen zur Erfassung von Angststörungen im Kindes- und Jugendalter. *Klinische Diagnostik und Evaluation*, 3, 142–163.
- Möller, J. & Trautwein, U. (2009). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 179–203). Berlin, Heidelberg: Springer.
- Moosbrugger, H. & Schermelleh-Engel, K. (2012). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2., akt. und überarb. Aufl., S. 325–343). Berlin, Heidelberg: Springer.
- Muris, P. & Field, A. P. (2011). The “normal” development of fear. In A. P. Field & W. K. Silverman (Eds.), *Anxiety Disorders in Children and Adolescents* (2nd ed., pp. 76–89). Cambridge: Cambridge University Press.
- Petermann, U. & Petermann, F. (2015). *Training mit sozial unsicheren Kindern. Behandlung von sozialer Angst, Trennungsangst und generalisierter Angst*. Mit Arbeitsmaterial (11., überarb. und erw. Aufl.). Weinheim: Beltz.
- R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing [Computer software]. Viena, Austria: R Foundation for Statistical Computing. Verfügbar unter: <https://www.R-project.org/>
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47, 667–696. <https://doi.org/10.1080/00273171.2012.715555>
- Riley, A. W. (2004). Evidence that school-age children can self-report on their health. *Ambulatory Pediatrics*, 4, 371–376. <https://doi.org/10.1367/A03-178R.1>
- Rosseel, Y. (2012). lavaan. An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48 (2), 1–36 <https://doi.org/10.18637/jss.v048.i02>
- Rubin, K. H., Chen, X. & Hymel, S. (1993). Socioemotional characteristics of withdrawn and aggressive children. *Merrill-Palmer Quarterly*, 39, 518–534.
- Saile, H. & Kison, K. (2002). Erfassung sozialer Unsicherheit bei Kindern. *Diagnostica*, 48, 6–11. <https://doi.org/10.1026/0012-1924.48.1.6>
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In D.D.H. Heijmans, D.S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis: A Festschrift for Heinz Neudecker* (pp. 233–247). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Scharfstein, L. A. & Beidel, D. C. (2015). Social skills and social acceptance in children with anxiety disorders. *Journal of Clinical Child and Adolescent Psychology*, 44, 826–838. <https://doi.org/10.1080/15374416.2014.895938>
- Schmidtendorf, S., Wiedau, S., Asbrand, J., Tuschen-Caffier, B. & Heinrichs, N. (2018). Attentional bias in children with social anxiety disorder. *Cognitive Therapy and Research*, 42, 273–288. <https://doi.org/10.1007/s10608-017-9880-7>
- Schneider, S. (2004). Entwicklungspathologische Grundlagen. In S. Schneider (Hrsg.), *Angststörungen bei Kindern und Jugendlichen: Grundlagen und Behandlung* (S. 3–16). Berlin: Springer.
- Schneider, S. & Döpfner, M. (2004). Leitlinien zur Diagnostik und Psychotherapie von Angst- und Phobischen Störungen im Kindes- und Jugendalter: Ein evidenzbasierter Diskussionsvorschlag. *Kindheit und Entwicklung*, 13, 80–96. <https://doi.org/10.1026/0942-5403.13.2.80>
- Schreier, S.-S. & Heinrichs, N. (2008). Die Elternversion der deutschen Fassung der Social Anxiety Scale for Children Revised (SASC-R-D): Notwendig, hinreichend oder überflüssig? *Klinische Diagnostik und Evaluation*, 1, 430–446.
- Schweizer, K. (2011). On the changing role of cronbach's α in the evaluation of the quality of a measure. *European Journal of Psychological Assessment*, 27, 143–144. <https://doi.org/10.1027/1015-5759/a000069>
- Stewart, S. L. & Rubin, K. H. (1995). The social problem-solving skills of anxious-withdrawn children. *Development and Psychopathology*, 7, 323–336. <https://doi.org/10.1017/S0954579400006532>
- Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines:

- An illustration using Mplus and the lavaan/semtools packages. *Structural Equation Modeling: A Multidisciplinary Journal*, 27, 111–130. <https://doi.org/10.1080/10705511.2019.1602776>
- Thiels, C. & Schmitz, G. S. (2008). Selbst- und Fremdbeurteilung von Verhaltensauffälligkeiten bei Kindern und Jugendlichen. *Kindheit und Entwicklung*, 17, 118–125. <https://doi.org/10.1026/0942-5403.17.2.118>
- Verduin, T. L. & Kendall, P. C. (2008). Peer perceptions and liking of children with anxiety disorders. *Journal of Abnormal Child Psychology*, 36, 459–469. <https://doi.org/10.1007/s10802-007-9192-6>
- World Health Organization. (2004). *International statistical classification of diseases and related health problems* (10. rev., 2. Ed.). Geneva.

Förderung

Open Access-Veröffentlichung ermöglicht durch die Bergische Universität Wuppertal

Philipp Nicolay

Institut für Bildungsforschung
Bergische Universität Wuppertal
Gaußstraße 20
42119 Wuppertal
nicolay@uni-wuppertal.de