

Konstruktion und Validierung einer Kurzform der Skala Angst vor negativer Bewertung (SANB-5)

Christoph J. Kemper¹, Johannes Lutz² und Jürgen Neuser³

Zusammenfassung

Das Persönlichkeitsmerkmal Bewertungsängstlichkeit bezieht sich auf die Neigung, in sozialen Situationen eine negative Bewertung durch Interaktionspartner zu fürchten. Im deutschen Sprachraum wird Bewertungsängstlichkeit oft mit der *Skala Angst vor negativer Bewertung* (SANB, Vormbrock & Neuser, 1983) erfasst. Die vorliegende Untersuchung beschreibt die Konstruktion und Validierung einer um 75 % gekürzten Version dieser Skala (SANB-5). Die Überprüfung der Testgüte erfolgte an einer klinischen und zwei nicht-klinischen Stichproben (insgesamt: $N = 2929$). Die interne Konsistenz der SANB-5 liegt bei Cronbach $\alpha > .84$. Das einfaktorielle Testmodell beschreibt die Daten der Stichproben angemessen. Zudem erwies sich das einfaktorielle Modell als messinvariant im Hinblick auf den Vorgabemodus (Papier vs. Online). Differentielle Korrelationen zu den Big-5, Ängstlichkeit, Angstsensitivität und Depressivität werden berichtet. Ein Vergleich der Testgütekriterien von SANB-5 und SANB zeigt, dass sich trotz der deutlichen Reduktion des ursprünglichen Itemsatzes um 75 %, keine substantiellen Einbußen im Hinblick auf Messgenauigkeit und Konstruktvalidität ergeben.

Schlüsselwörter

Bewertungsängstlichkeit – Bewertungsangst – Soziale Phobie – Soziale Angst – Angst vor negativer Bewertung

Abstract

Development and validation of a short version of the Fear of Negative Evaluation Scale (SANB-5)

Fear of negative evaluation is a personality trait denoting the tendency to dread being evaluated unfavorably by others. In German-speaking countries, fear of negative evaluation is usually measured by a German adaptation (SANB) of the *Fear of Negative Evaluation Scale* (FNE, Watson & Friend, 1969). The present research describes the construction and validation of a 5-item short version (SANB-5) of the German adaptation.

1 GESIS – Leibniz-Institut für Sozialwissenschaften, Mannheim

2 Universität Potsdam

3 Institut für medizinische und pharmazeutische Prüfungsfragen, Mainz

Clinical and non-clinical samples were used (overall $N = 2929$). Internal consistencies of Cronbach $\alpha > .84$ were achieved. In all samples an onefactorial model adequately fitted the data. In addition, measurement (scalar) invariance was supported concerning presentation mode of the items (paper-and-pencil vs. online). Furthermore, differential correlations with the Big Five personality factors, anxiety, anxiety sensitivity, and depressive symptoms are reported. Despite considerable reductions of the original item set by 75%, no substantial decrease in reliability or construct validity was observed.

Keywords

Fear of negative evaluation – evaluation anxiety – social phobia – social fear – FNE

Einleitung

Das Konstrukt Bewertungsängstlichkeit

Die Angst vor negativer Bewertung wird von Watson und Friend (1969) beschrieben als Befürchtungen über mögliche negative Bewertungen durch andere, Belastung durch diese negativen Bewertungen, Vermeidung von Bewertungssituationen und die Erwartung, dass andere einen in der Zukunft negativ bewerten werden. Wenn dieses Verhalten bei einer Person wiederholt, also über einen längeren Zeitraum, und in verschiedenen alltäglichen Situationen gezeigt wird, kann von einer Neigung zur Bewertungsangst gesprochen werden. Anders als Test- bzw. Prüfungsängstlichkeit bezieht sich Bewertungsängstlichkeit nicht nur auf klar definierte Bewertungssituationen (Prüfung), sondern auf alle sozialen Situationen.

Skalen zur Erfassung von Bewertungsängstlichkeit werden zum Beispiel in der klinisch-psychologischen-, der differentiellen- oder der sozialpsychologischen Forschung eingesetzt (Gilbert & Meyer, 2005a 2005b; Hinrichsen, Waller & Dhokia, 2007; Oaten, Williams, Jones, & Zadro, 2008; Zadro, Boland & Richardson, 2006). Um Fortschritte in diesen Forschungsbereichen zu erzielen, sind psychometrische Verfahren notwendig, die eine reliable und valide Erfassung des Konstrukts erlauben. Im deutschen Sprachraum wird Bewertungsängstlichkeit mit einer Übersetzung der Originalskala (FNE, *Fear of Negative Evaluation-Scale*; Watson & Friend, 1969), der *Skala Angst vor negativer Bewertung* (SANB; Vormbrock & Neuser, 1983), operationalisiert. Für den englischen Sprachraum existieren mittlerweile zwei überarbeitete Versionen der Originalskala. Ziel der vorliegenden Untersuchung war die Überarbeitung der deutschen Übersetzung.

Im Folgenden soll das Konstrukt Bewertungsängstlichkeit zunächst auf der konzeptuellen Ebene betrachtet und von anderen Persönlichkeitsmerkmalen abgegrenzt werden. Anschließend werden einige empirische Ergebnisse zur Beziehung von Bewertungsängstlichkeit und anderen Konstrukten vorgestellt.

Maße zur Erfassung von Bewertungsängstlichkeit werden oftmals als Maße der sozialen Ängstlichkeit (z. B. Hinrichsen, Waller & Dhokia, 2007; Winton, Clark & Edelmann, 1995) oder auch der sozialen Phobie (z. B. Duke, Krishnan, Faith & Storch, 2006) verwendet. Nach Weeks, Heimberg, Fresco, Hart und Turk (2005) sind Bewertungsängstlichkeit und soziale Ängstlichkeit allerdings nicht gleichzusetzen. Sie sehen Bewertungsängstlichkeit als die zentrale Komponente der sozialen Ängstlichkeit an. Neuser (2003) bezeichnet Bewertungsängstlichkeit als die kognitive Komponente der sozialen Ängstlichkeit (im Sinne der Einteilung von Liebert und Morris, 1967). Reiss (1991) ordnet Bewertungsängstlichkeit neben Angstsensitivität und Verletzungssensitivität im Rahmen seines Erwartungsmodells als eine der drei grundlegenden Arten von Furcht (Sensitivitäten) ein, auf die alle anderen Arten der Furcht zurückgeführt werden können.

Eine weitergehende Einordnung des Konstrukts Bewertungsängstlichkeit erlauben die zahlreichen empirischen Befunde, die die Relevanz des Konstrukts für verschiedene Bereiche der psychologischen Forschung unterstreichen. Im Rahmen des Fünffaktorenmodells der Persönlichkeit (Costa & McCrae, 1992; Goldberg, 1993) lässt sich Bewertungsängstlichkeit klar dem Faktor Neurotizismus zuordnen (vgl. Vormbrock & Neuser, 1983). Mit Maßen der sozialen Ängstlichkeit weist Bewertungsängstlichkeit üblicherweise mittlere Korrelationen auf (Leary, 1983; Vormbrock & Neuser, 1983; Weeks et al., 2005; Carleton, Collimore & Asmundson, 2007). Weiterhin finden sich substantielle Zusammenhänge mit allgemeiner Ängstlichkeit (Vormbrock & Neuser, 1983) und Angstsensitivität (Kemper, Ziegler & Taylor, 2009).

Auch bei der Selbstregulation spielt Bewertungsängstlichkeit offenbar eine Rolle. So zeigen Oaten et al. (2008) und Zadro et al. (2005), dass Bewertungsängstlichkeit den Effekt von sozialer Ausgrenzung auf die Selbstregulation moderiert. Probanden mit hoher Bewertungsängstlichkeit empfanden eine experimentell induzierte Ausgrenzung als schwerwiegender und dauerhafter im Vergleich zu Probanden mit niedriger Bewertungsängstlichkeit. Darüber hinaus zeigten sie in Folge der Ausgrenzung Defizite in der Selbstregulation, wie z. B. exzessives Essen oder vermehrtes Essen ungesunder Speisen. In diese sozialpsychologischen Befunde lassen sich Befunde aus der klinisch-psychologischen Forschung zu Essstörungen einreihen. Nach diesen scheinen Personen mit hoher im Gegensatz zu Personen mit geringer Bewertungsängstlichkeit zu restriktivem Essverhalten zu neigen (Hinrichsen et al., 2007; Gilbert & Meyer, 2005a, 2005b).

Die Messung von Bewertungsängstlichkeit

Die *Fear of Negative Evaluation Scale* (FNE) von Watson und Friend (1969) soll das Konstrukt Bewertungsängstlichkeit abbilden. Laut der Autoren handelt es sich bei Bewertungsängstlichkeit um ein eindimensionales Konstrukt. Die FNE-Skala besteht aus 30 dichotomen Items, von denen 17 negativ kodiert

sind (interne Konsistenz nach Kuder-Richardson-Statistik: .94 – .96). Untersuchungen zur Konstruktvalidität zeigten, dass die FNE mittel bis hoch mit allgemeiner Ängstlichkeit und stärker mit dem Bedürfnis nach Anerkennung korreliert (Watson & Friend, 1969). Ein negativer Zusammenhang zeigte sich mit dominantem Verhalten. Bis heute wurden zwei Revisionen der FNE publiziert, die BFNE (*Brief Fear of Negative Evaluation-Scale*; Leary, 1983) und die BFNE-II (*Brief Fear of Negative Evaluation Scale-Revised*; Carleton, McCreary, Norton, & Asmundson, 2007).

Anlass für die erste Revision (BFNE) war die Länge der FNE-Skala (Weeks et al., 2005; Leary, 1983). Leary (1983) kürzte die Skala auf 12 Items, von denen vier negativ kodiert sind, und führte eine fünfstufige Antwortskala ein. Die interne Konsistenz wird mit Cronbach's $\alpha = .90$ angegeben. Bei faktorenanalytischen Untersuchungen der BFNE zeigte sich in mehreren Studien eine zweifaktorielle Struktur und nicht die von Watson und Friend postulierte einfaktorielle Struktur (Carleton, et al., 2006; Duke et al., 2006; Weeks et al., 2005; Rodebaugh et al., 2004). In allen Studien wurde jeweils ein zweiter atheoretischer Faktor identifiziert, der durch die vier negativ kodierten Items gebildet wurde (Faktorinterkorrelation: $-.42$ bis $-.31$). Ein solcher Faktor muss keine inhaltliche Bedeutung haben (Spector, van Katwyk, Brannick & Chen, 1997). Er kann beispielsweise Resultat eines unterschiedlichen Antwortverhaltens der Probanden auf die im Sinne des Konstrukts negativ formulierten Items im Gegensatz zu den positiv formulierten Items sein. Rodebaugh et al. (2004) und Weeks et al. (2005) nahmen an, dass die negative Itemformulierung zu einer Verwirrung der Probanden führt. In der Folge könnten die Probanden, abhängig von der Richtung der Itemformulierung, unterschiedliche Konzepte des zu erfassenden Konstrukts generieren.

Aufgrund der problematischen faktoriellen Validität und der möglichen Effekte der konstruktnegativ formulierten Items, kam es zu einer Modifikation der BFNE durch Carleton et al. (2006; 2007). Dabei wurden die konstrukt-negativen Items umformuliert (Carleton et al., 2006) und vier Items eliminiert (Carleton et al., 2007). Es resultierte die BFNE-II mit acht ausschließlich konstruktpositiv formulierten Items und fünfstufigem Antwortformat. Bei einer Überprüfung mittels CFA zeigte das einfaktorielle Modell eine exakte Passung auf die Daten. Die interne Konsistenz der BFNE-II wird mit $\alpha = .96$ angegeben. Sowohl die Acht- als auch die Zwölf-Item-Version zeigten gleichwertige moderate bis hohe positive Korrelationen mit Maßen der sozialen Ängstlichkeit. Korrelationen mit den Facetten der Angstsensitivität waren positiv und lagen im mittleren Bereich. Der größte Zusammenhang bestand dabei mit *AS Bedenken sozial*. Geringere positive Korrelationen fanden sich für beide BFNE-II-Versionen und Krankheits- und Verletzungsangst.

Eine deutsche Übersetzung der FNE ist die Skala *Angst vor negativer Bewertung* (SANB; Vormbrock & Neuser, 1983). Im Zuge der Konstruktion der deutschen Fassung, führten die Autoren eine exploratorische Faktorenanalyse (EFA) der 30 übersetzten FNE-Items durch. Der Scree-Test legte eine zweifak-

torielle Struktur nahe. Da auf dem zweiten Faktor allerdings nur vier Items substanzielle Ladungen aufwiesen, wurde dieses Modell zugunsten des besser interpretierbaren einfaktoriellen Modells verworfen. Aus den 30 übersetzten Items der FNE wurden 20 ausgewählt, von denen sechs negativ kodiert waren. Anstelle des dichotomen Antwortformates wurde eine vierstufige Antwortskala gewählt. Die Trennschärfen der 20 selektierten SANB-Items waren hoch ($.45 < r_{it} < .74$), die interne Konsistenz ebenfalls ($\alpha = .92$). Im Rahmen der Konstruktvalidierung ergaben sich folgende Ergebnisse: Die SANB korrelierte substanziell mit dem State-Trait-Angstinventar (STAI; Laux, Glanzmann, Schaffner, & Spielberger, 1981), Maßen sozialer Ängstlichkeit und der Persönlichkeitsdimension Neurotizismus. Eine überarbeitete Fassung der SANB, die der in der Literatur geäußerten Kritik an den englischen Skalen FNE und BFNE Rechnung trägt, gibt es bisher nicht.

Ziel der Untersuchung

Ziel der vorliegenden Untersuchung war es, eine revidierte ökonomische Fassung der SANB mit hinreichender psychometrischer Güte für den Einsatz in der Forschung zu erstellen, die den in der Literatur geäußerten Kritikpunkten Rechnung trägt. Es sollte (a) eine stark verkürzte Version der SANB erstellt werden. Die sehr hohen internen Konsistenzen der englischen Skalen und der SANB legten einen Spielraum für Kürzungen nahe, ohne die Breite der Konstruktooperationalisierung wesentlich einschränken zu müssen. Weiterhin sollte (b) die faktorielle Validität der Operationalisierung verbessert werden. Die modifizierte Skala sollte im Gegensatz zu den Befunden verschiedener Studien, die eine zweifaktorielle Struktur mit einem atheoretischen zweiten Faktor fanden (z. B. Weeks et al., 2005; Rodebaugh et al., 2004), eine einfaktorielle Struktur aufweisen. Um die Generalisierbarkeit der einfaktoriellen Struktur zu erhöhen, wurde zur Untersuchung der faktoriellen Validität neben zwei nicht-klinischen Stichproben auch eine klinische Stichprobe herangezogen. Schließlich sollte (c) gezeigt werden, dass die Kürzung der Skala deren psychometrische Güte nur unwesentlich einschränkt. Neben der Reliabilität wurden konvergente und diskriminante Validitätskoeffizienten ermittelt. Eine hohe Korrelation der gekürzten Skala wurde mit der Traitskala des *State-Trait-Angstinventars* erwartet, da diese insbesondere für Angstreaktionen in sozialen Situationen prädiktiv sein soll. Endler (1975) konnte zeigen, dass anhand der Traitskala Werte auf der Stateskala des STAI nur für Situationen vorhergesagt werden können, in denen die Gefahr besteht, das Selbstwertgefühl zu verletzen oder sich zu blamieren (soziale Situationen), aber nicht für physisch bedrohliche Situationen. Es wurde weiterhin erwartet, dass die gekürzte Skala hoch mit der Skala *Bedenken sozial* aus dem Ängstsensitivitätsindex-3 (ASI-3; Kemper et al., 2009) korreliert. Menschen mit hoher Bewertungsängstlichkeit fürchten negative Bewertung durch Andere. Menschen mit hoher Ausprägung auf dem *AS-Faktor Bedenken sozial* fürchten, dass die

Angstsymptome, für die sie sensitiv sind, von anderen entdeckt und negativ bewertet werden könnten (Taylor, 1999). Eine konzeptuelle Nähe von Bewertungsängstlichkeit und *AS-Bedenken sozial* ist daher gegeben. Deutlich niedriger sollten hingegen die Korrelationen der gekürzten Skala mit den anderen beiden Skalen des ASI-3 ausfallen. Im Hinblick auf die fünf Hauptfaktoren der Persönlichkeit nach dem Fünffaktorenmodell wurde eine hohe Korrelation mit der Skala *Neurotizismus* des *Big-Five-Inventory* (BFI; Rammstedt, 1997) erwartet und nur geringe Korrelationen mit *BFI-Verträglichkeit*, *BFI-Gewissenhaftigkeit* und *BFI-Offenheit*, da Bewertungsangst zu den negativen Affekten zählt, die im Neurotizismuskonstrukt thematisiert werden. Außerdem wurden aufgrund der konzeptuellen und empirischen Überlappung der Konstrukte Ängstlichkeit und Depressivität eine substantielle Korrelation der gekürzten Skala mit der *Allgemeinen Depressions Skala* (ADS-K; Hautzinger & Bailer, 1993) erwartet.

Methoden

Studiendesign und Vorgehen.

Die Analysen wurden an zwei nicht-klinischen Stichproben (fortan: Online- bzw. Paper-Pencil-Stichprobe) und einer klinischen Stichprobe durchgeführt, die im Rahmen anderer Studien des Erstautors erhoben wurden. Die Onlinestichprobe wurde größtenteils über die Email-Verteiler von Psychologie-Fachschaften verschiedener deutscher Universitäten rekrutiert (Erhebungszeitraum: III./IV. Quartal 2007), während die Paper-Pencil-Stichprobe im Rahmen von Lehrveranstaltungen am psychologischen Institut der Universität Mainz erhoben wurde (Erhebungszeitraum: IV. Quartal 2007). Bei der klinischen Stichprobe handelt es sich um eine naturalistische Stichprobe aus der psychotherapeutischen Routineversorgung (Erhebungszeitraum: IV. Quartal 2007). Die Patienten wurden von ihren Therapeuten nach ihrer Bereitschaft, an einem Forschungsprojekt der Universität Mainz mitzuwirken, befragt und füllten die Fragebögen zwischen zwei Therapiesitzungen zu Hause aus. In alle drei Stichproben wurden nur Personen aufgenommen, die zum Zeitpunkt der Erhebung 18 Jahre oder älter waren.

Messinstrumente.

Es wurden insgesamt fünf Selbstbeurteilungsinstrumente eingesetzt. Alle Probanden bekamen die ursprüngliche 20-Item-Version der SANB vorgelegt. Zur Überprüfung der konvergenten und diskriminanten Validität füllten die Probanden aller Stichproben zusätzlich den ASI-3 mit Skalen für globale, somatische, soziale und kognitive Angstsensitivität und die Kurzversion der ADS aus. Desweiteren wurde in der klinischen Stichprobe das STAI und

in der Paper-Pencil-Stichprobe das Big-Five-Inventory mit Skalen für Neurotizismus, Extraversion, Offenheit, Verträglichkeit und Gewissenhaftigkeit bearbeitet.

Statistische Analysen.

Die Itemselektion und Validierung der Kurzskaala wurde in vier Schritten durchgeführt. Zur Überprüfung der faktoriellen Validität wurden konfirmatorische Faktoranalysen mit AMOS 7 und MPlus 5.1 durchgeführt. Bei allen übrigen Analysen kam SPSS 15.0 zum Einsatz. (1) Die Itemselektion wurde aufgrund der durchweg hohen Trennschärfen aller SANB-Items ausschließlich an inhaltlichen Kriterien ausgerichtet (Maximierung der Inhaltsvalidität). Zunächst wurden die 20 Items der SANB von einer Expertenrunde⁴ überprüft. Jedes Testitem wurde auf Grundlage des Fragebewertungssystems (Faulbaum, Prüfer, & Rexroth, 2009), einem systematischen Ansatz zur Beurteilung der Qualität von Fragebogenitems, bewertet. Anschließend wurden Items für die Kurzversion gewählt, die keine im Fragebewertungssystem genannten problematischen Aspekte, wie beispielsweise unklare Bezüge, vage Bedeutung von Begriffen oder zu hohe Anforderungen an die kognitive Leistungsfähigkeit (vgl. Faulbaum et al., 2009) aufwiesen. Weiterhin wurden bei der Itemselektion Redundanzen, wie in der SANB vermieden, und darauf geachtet, dass die inhaltliche Breite des Konstrukts erhalten bleibt (vgl. z. B. Stanton, Sinar, Balzer & Smith, 2002). Fünf Items wurden in die Kurzversion (fortan SANB-5) übernommen.

(2) Danach wurden, getrennt für Online-, Paper-Pencil- und klinische Stichprobe, Itemstatistiken berechnet und die Reliabilität (Cronbach α) bestimmt.

(3) Im dritten Schritt wurde die faktorielle Validität der SANB-5 mit Hilfe von CFAs (ML-Schätzung) gesichert. Zunächst wurde die Passung des einfaktorischen Modells auf die Daten der nicht-klinischen Stichproben überprüft. Für die Beurteilung der Modellpassung aller CFA-Lösungen wurden, neben dem χ^2 -Test auf exakte Modellpassung, der Standardized Root Mean Square Residual (SRMR, sollte $\leq .08$ sein), der Comparative Fit Index (CFI, sollte $\geq .95$ sein) und der Root Mean Square Error of Approximation (RMSEA, sollte $\leq .06$ sein) berücksichtigt (Hu & Bentler, 1999; Beauducel & Wittmann, 2005). Anschließend wurde eine Invarianzprüfung durchgeführt, um feststellen zu können, ob das Testmodell für Personen, die den Test online bzw. als Papierversion bearbeiten, gleichermaßen Gültigkeit hat. Bei einer Invarianzprüfung werden Modellparameter simultan für zwei oder mehr Gruppen geschätzt (Multigruppen-Modell). Bestimmte Modellparameter werden sukzessive über die Stichproben restringiert. Dies erlaubt die Prüfung verschiedener Aspekte

4 Das Expertenteam setzte sich aus vier Experten für Angstforschung und Fragebogenkonstruktion zusammen (inklusive zweier Autoren).

der Invarianz (vgl. Vandenberg & Lance, 2000): konfigurale Invarianz (Gleichheit des Musters freier und fixierter Parameter über die Stichproben), metrische Invarianz (Gleichheit der Faktorladungen der Items), skalare Invarianz (Gleichheit der Regressionskonstanten der Items). Führt die Einführung von Restriktionen zu einer Veränderung des CFI gegenüber dem Modell der konfiguralen Invarianz um .002 oder mehr, ist von einer signifikanten Modellverschlechterung und somit einem Mangel an Invarianz im Bezug auf die zuvor restringierten Modellparameter auszugehen (Meade, Johnson & Braddy, 2008). Metrische Invarianz gilt als Voraussetzung für den Vergleich von Korrelationen oder Regressionskoeffizienten zwischen Stichproben, skalare Invarianz für den Vergleich von Mittelwerten (Vandenberg & Lance, 2000). Die klinische Stichprobe konnte aufgrund der geringen Stichprobengröße nicht in die Invarianzprüfung einbezogen werden. Die Daten dieser Stichprobe wurden in MPLUS analysiert (CFA-Modell, ML-Schätzung, TYPE = COMPLEX). Dabei wurde die Diagnosegruppe als Clustervariable verwendet.

(4) Im vierten Schritt wurden Pearson-Korrelationen mit verschiedenen Selbstbeurteilungsinstrumenten berechnet, um Hinweise zur konvergenten und diskriminanten Validität der SANB-5 zu erhalten. Es wurde ein zweistufiges Vorgehen gewählt, damit belegt werden konnte, dass durch die Kürzung lediglich Redundanzen im Itemmaterial beseitigt wurden und die Beziehungen im nomologischen Netzwerk weitgehend unberührt bleiben. Erstens wurde geprüft, ob sich die Validitätskoeffizienten von SANB-5 und SANB signifikant unterscheiden. Zweitens wurden Partialkorrelationen zwischen dem Summenwert der 15 eliminierten SANB-Items und den Validierungsmaßen unter Kontrolle der SANB-5 berechnet. Damit sollte weiter untermauert werden, dass die eliminierten Items keinen inkrementellen Beitrag über die SANB-5 hinaus für die Prädiktion der Validierungsmaße haben.

Sämtlichen Signifikanztests der Untersuchung lag das konventionelle α -Niveau von .05 zugrunde.

Ergebnisse

Stichproben

Die Onlinestichprobe bestand aus $N = 2603$ Probanden zwischen 16 und 81 Jahren ($M = 24.83$, $SD = 6.35$), mehrheitlich Studierende deutscher Hochschulen (88.1%) weiblichen Geschlechts (68.6%). Die Paper-Pencil-Stichprobe bestand aus $N = 256$ Probanden in einem Alter von 18 bis 58 Jahren ($M = 23.88$, $SD = 7.26$). Die Probanden waren überwiegend Studierende (92.6%) und weiblichen Geschlechts (78.9%). Die klinische Stichprobe setzte sich aus $N = 75$ Patienten niedergelassener Psychotherapeuten aus der Region Mainz zusammen. Die Patienten waren zwischen 16 und 63 Jahren alt ($M = 39.13$, $SD = 12.13$) und zu 84% weiblichen Geschlechts. Bei allen wurde durch die Therapeuten

die Hauptdiagnose Affektive Störung (F30-F39 nach ICD-10) oder Neurotische, Belastungs- oder somatoforme Störung (F40-F48) vergeben.⁵

Itemstatistiken und Reliabilität. Auf den Items der SANB-5 gaben die Probanden der beiden nicht-klinischen Stichproben im Mittel ca. eine „2“ an (siehe Tabelle 1). Diese Ausprägung entspricht der Antwortkategorie „trifft manchmal zu“. Die Itemmittelwerte liegen unterhalb des Skalenmittelpunkts von 2.5. Die Verteilungen der Items sind daher geringfügig linkssteil. Im Vergleich dazu sind die Items in der klinischen Stichprobe annähernd normal verteilt. Die Mittelwerte der Items liegen nahe am Skalenmittelpunkt oder geringfügig oberhalb des Skalenmittelpunkts zwischen 2.5 und 3 („3“ entspricht der Antwortkategorie „trifft oft zu“). Während die Streuungen der Items in den beiden nicht-klinischen Stichproben fast identisch sind, weisen die Items in der klinischen Stichprobe deskriptiv eine höhere Streuung auf. Die Reliabilität des Summenwerts der SANB-5 liegt bei $.84 < \alpha < .94$ (siehe Tabelle 3) und kann somit als moderat bis hoch angesehen werden (vgl. Bühner, 2006).

Tabelle 1: Itemstatistiken der SANB-5 getrennt nach Stichproben

	Onlinestichprobe (N = 2603)				Paper-Pencil-Stichprobe (N = 256)				Klinische Stichprobe (N = 75)			
	M	SD	Sf	Kurt	M	SD	Sf	Kurt	M	SD	Sf	Kurt
1: Ich fürchte, Falsches zu tun oder zu sagen.	2.28	.92	.34	-.68	2.17	.94	.38	-.74	2.71	.99	-.22	-.99
2: Ich bin beunruhigt darüber, welchen Eindruck ich auf jemanden mache.	2.30	.90	.26	-.69	2.36	.90	.22	-.68	2.52	.95	.14	-.89
3: Ich fürchte, dass andere mich nicht anerkennen.	2.17	.98	.37	-.92	2.09	.96	.49	-.74	2.71	.99	-.22	-.99
4: Wenn ich mich mit jemandem unterhalte, habe ich Angst davor, was der andere von mir denkt.	1.91	.88	.67	-.38	1.87	.88	.70	-.38	2.39	1.03	.08	-1.12
5: Ich habe Angst vor Leuten, die meine Schwächen bemerken.	2.05	.91	.51	-.57	1.98	.88	.60	-.37	2.55	1.02	.15	-1.12
SANB-5 Summenwert	10.71	3.59	.47	-.42	10.46	3.68	.62	-.25	12.87	4.29	.09	-.94
SANB Summenwert	47.53	11.56	.23	-.47	48.27	11.74	.20	-.43	54.20	12.48	-.16	-.26

Anmerkungen: M = Mittelwert, SD = Standardabweichung, Sf = Schiefe, Kurt = Kurtosis, Antwortkategorien der SANB-5: trifft fast nie zu (1), trifft manchmal zu (2), trifft oft zu (3), trifft fast immer zu (4)

5 Genauere Informationen über die Diagnosestellung lagen nicht vor. Allerdings wurden die Patienten ausschließlich von niedergelassenen Psychotherapeuten mit mehrjähriger Berufserfahrung diagnostiziert.

Faktorielle Validität. Die faktorielle Validität der SANB-5 wurde durch die Anpassung der Grundmodelle und der Multigruppenmodelle geprüft. Zunächst wurde das Testmodell auf die Daten aller Stichproben separat angepasst (Grundmodelle). Anschließend wurde das Testmodell auf die beiden nicht-klinischen Stichproben simultan angepasst und Parameterrestriktionen eingeführt, um die Invarianz des Testmodells im Hinblick auf den Erhebungsmodus zu belegen (Multigruppenmodelle). Tabelle 2 zeigt die Ergebnisse der CFAs zur faktoriellen Validität der SANB-5. Das Testmodell zeigt in allen drei Stichproben eine gute Passung auf die Daten. Alle fünf Items weisen substantielle Ladungen auf dem latenten Faktor auf (siehe Abb. 1). Zwischen den Items der SANB-5 traten keine substantiellen Residualkorrelationen auf. Der χ^2 -Test auf exakte Modellpassung weist für die verwendeten Stichproben, mit Ausnahme der Onlinestichprobe, auf eine exakte Modellpassung hin. Die approximativen Fitindizes zeigen jedoch für alle drei Stichproben, auch die Onlinestichprobe, eine gute Passung des Testmodells auf die Daten an.

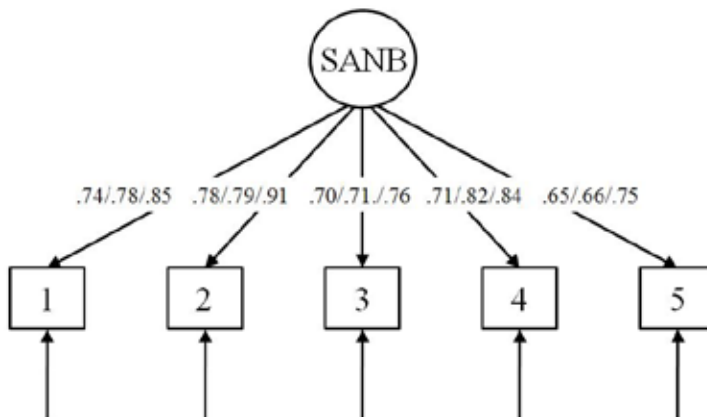


Abbildung 1: CFA-Modell der SANB-5 mit Faktorladungen für die Online-, Paper-Pencil- und klinische Stichprobe.

Die faktorielle Validität der SANB-5 wird auch durch die Multigruppenmodelle für die beiden nicht-klinischen Stichproben gestützt. Wie aus Tabelle 2 zu entnehmen ist, lag für die SANB-5 konfigurale Invarianz vor. Demnach wurde die latente Variable Bewertungsängstlichkeit in beiden nicht-klinischen Stichproben von denselben Items indiziert. Ein Vergleich des Grundmodells (konfigurale Invarianz) mit einem Modell, in dem Faktorladungen und Regressionskonstanten über die Stichproben gleich gesetzt wurden (Modell skalare Invarianz), zeigt keine substantielle Veränderung im CFI. Demnach gilt für die SANB-5 im Hinblick auf die in der Untersuchung realisierten Erhebungsmodi skalare Invarianz.

Tabelle 2: Ergebnisse der CFAs zur faktoriellen Validität der SANB-5

	$\chi^2(df)$	p	SRMR	RMSEA (90 % CI)	CFI
<u>Grundmodelle</u>					
Onlinestichprobe ¹	25.47 (5)	< .001	.013	.040 (.025 – .056)	.9956
Paper-Pencil-Stichprobe ²	3.24 (5)	.662	.012	.000 (.000 – .069)	1.00
Klinische Stichprobe ³	4.63 (5)	.460	.019	.000 (.000 – .071)	1.00
<u>Multigruppenmodelle</u>					
konfigurale Invarianz	28.72 (10)	< .01	.013	.026 (.015 – .037)	.9964
metrische Invarianz	33.73 (14)	< .01	.013	.022 (.012 – .032)	.9964
skalare Invarianz	45.19 (19)	< .01	.013	.023 (.015 – .031)	.9947

Anmerkung: ¹N = 2603; ²N = 256; ³N = 75.

Konvergente und diskriminante Validität. In Tabelle 3 finden sich die Ergebnisse zur Prüfung der konvergenten und diskriminanten Validität getrennt für die drei Stichproben. Die SANB-5 korreliert sehr hoch mit der Originalversion SANB ($.90 < r < .92$) und mit den 15 im Zuge der Modifikation eliminierten Items ($.80 < r < .83$). Im Hinblick auf die fünf Hauptfaktoren der Persönlichkeit ist die SANB-5 stark mit *BFI-Neurotizismus* assoziiert ($r = .66$). Es findet sich auch eine, allerdings deutlich geringere, moderat negative Korrelation mit *BFI-Extraversion* ($r = -.30$). Mit den übrigen Faktoren des BFI ist die SANB-5 im Wesentlichen unkorreliert (diskriminante Validität).

Weiterhin korreliert die SANB-5 in allen Stichproben hoch mit ADS-K und ASI-3, sowie mit dem in der klinischen Stichprobe eingesetzten STAI. Die Korrelation der SANB-5 mit dem ASI-3 geht offenbar hauptsächlich auf *ASI3-Bedenken sozial* zurück. Mit *ASI3-Bedenken kognitiv* sind die Korrelationen signifikant niedriger, mit *ASI3-Bedenken somatisch* am niedrigsten (jeweils $p < .05$). Dieses Korrelationsmuster wurde sowohl für die klinische als auch für die beiden nicht-klinischen Stichproben gefunden. Die SANB-5 korreliert ebenfalls substantiell und hoch mit der Traitskala des STAI ($r = .65$).

Der Vergleich der Validitätskoeffizienten zwischen der SANB-5 und der ursprünglichen Version mit 20 Items zeigt nur unwesentliche Unterschiede in der Validität. Sechs von einundzwanzig Koeffizienten haben zwar einen signifikanten p -Wert, die absolute Höhe der Unterschiede ist allerdings gering ($.02 < \Delta r < .06$). Eine Ausnahme bilden *BFI-Verträglichkeit* und *ASI3-Bedenken sozial*. Die Partialkorrelationen zwischen dem Summenwert der 15 eliminierten Items und den Validierungsmaßen sind mit Ausnahme der Onlinestichprobe und *AS-Bedenken sozial* in der Paper-Pencil-Stichprobe durchweg nicht signifikant. Die erklärte Varianz der Partialkorrelation liegt bei 0–5 %.

Tabelle 3: Korrelationen zwischen der SANB(-5) und den Validierungsmaßen sowie Partialkorrelationen zwischen dem Summenwert der eliminierten Items (SANB-15) und den Validierungsmaßen unter Kontrolle der SANB-5 in drei Stichproben.

	Cronbach α	SANB-5	SANB	p	SANB-15 SANB-5	r^2
<i>Onlinestichprobe (N = 2603)</i>						
SANB-5	.84					
SANB	.92	.90				
ADS-K	.89	.49	.48	$p > .05$.09**	.01
ASI3	.87	.59	.59	$p > .05$.20**	.04
ASI3-BSZ	.76	.59	.61	$p < .01$.23**	.05
ASI3-BKO	.79	.51	.49	$p < .01$.19**	.04
ASI3-BSM	.78	.32	.34	$p < .05$.13**	.02
<i>Paper-Pencil-Stichprobe (N = 256)</i>						
SANB-5	.87					
SANB	.93	.91**				
ADS-K	.89	.41**	.38**	$p > .05$.01	.00
ASI3	.87	.55**	.53**	$p > .05$.09	.01
ASI3-BSZ	.79	.55**	.56**	$p > .05$.17**	.03
ASI3-BKO	.80	.49**	.46**	$p > .05$.03	.00
ASI3-BSM	.76	.27**	.24**	$p > .05$	-.01	.00
BFI-O	.79	-.06	-.03	$p > .05$.05	.00
BFI-C	.79	-.08	-.02	$p < .05$.12	.01
BFI-E	.88	-.30*	-.28*	$p > .05$	-.02	.00
BFI-A	.73	-.19	-.09	$p < .01$.18	.03
BFI-N	.82	.66**	.65**	$p > .05$.16	.03
<i>Klinische Stichprobe (N = 75)</i>						
SANB-5	.94					
SANB	.93	.92**				
STAI	.93	.65**	.62**	$p > .05$.09	.01
ADS-K	.92	.47**	.42**	$p > .05$	-.04	.00
ASI3	.92	.69**	.64**	$p > .05$.01	.00
ASI3-BSZ	.86	.73**	.66**	$p < .05$	-.05	.00
ASI3-BKO	.89	.60**	.56**	$p > .05$.02	.00
ASI3-BSM	.88	.39**	.38**	$p > .05$.04	.00

Anmerkungen: SANB(-5) = Skala Angst vor negativer Bewertung(-5); SANB-15|SANB-5 = Partialkorrelation des Summenwerts der eliminierten Items der SANB und den Validierungskriterien unter Kontrolle der SANB-5; r^2 = erklärte Varianz der Partialkorrelation; ADS-K = Allgemeine Depressions Skala-Kurzform; ASI3-BSM = ASI-3 Skala Bedenken somatisch, ASI3-BSZ = ASI-3 Skala Bedenken sozial, ASI3-BKO = ASI-3 Skala Bedenken kognitiv, ASI-3 Gesamtwert; BFI-N = BFI-Neurotizismus, BFI-E = BFI-Extraversion, BFI-O = BFI-Offenheit für Erfahrung, BFI-V = BFI-Verträglichkeit, BFI-G = BFI-Gewissenhaftigkeit; * $p < .05$, ** $p < .01$, alle Differenzen der fettgedruckten Korrelationen $p < .05$ (jeweils zweiseitige Testung).

Diskussion

Die berichteten Befunde sprechen aus verschiedenen Gründen für eine adäquate psychometrische Güte der SANB-5 und legen deren Anwendung für Forschungszwecke nahe. Erstens wurde durch das Vorgehen bei der Itemselektion die Inhaltsvalidität der SANB-5 sichergestellt. Zweitens erwiesen sich die fünf selektierten Items als gute Indikatoren für das Konstrukt Bewertungsängstlichkeit. Das Testmodell ließ sich in klinischen und nicht-klinischen Stichproben bestätigen und zeigte sich zudem unabhängig vom Erhebungsmodus. Drittens konnte die konvergente und diskriminante Validität der SANB-5 über Zusammenhänge mit verschiedenen Selbstbeurteilungsinstrumenten belegt werden. Vor allem konnte gezeigt werden, dass die Kürzung nicht zu einer Veränderung der Beziehungen im nomologischen Netzwerk geführt haben. Viertens erwies sich die SANB-5 als hinreichend reliabel. Im Folgenden werden diese Befunde im Detail diskutiert.

Bei der Itemselektion wurde der Fokus auf die Inhaltsvalidität gelegt, da die Items der SANB bereits hohe Trennschärfen aufwiesen. Die Diskussion der Iteminhalte und Itemformulierungen im Expertengremium führte zu einer Auswahl von Items, die alle Aspekte der Bewertungsängstlichkeit aus der SANB enthalten, z. B. Angst davor, Fehler zu machen und Schwächen zu zeigen, Angst davor, mit den eigenen Aussagen, Meinungen oder Handlungen negativ aufzufallen und Angst davor, nicht anerkannt zu werden. Neben dem Erhalt der inhaltlichen Breite, wurde besonderes Augenmerk auf die Verständlichkeit der Itemformulierungen gelegt, da schwer verständliche oder missverständliche Testitems die Datenqualität mindern können. Mit dem Fragenbewertungssystem (Faulbaum et al., 2009) wurden systematisch problematische Testitems, z. B. solche mit unklaren Bezügen, mehrdeutigen Begriffen, umständlicher Syntax etc. identifiziert und eliminiert bis schließlich fünf inhaltlich und psychometrisch gute Items für die Kurzversion verblieben.

Im Gegensatz zu anderen Versionen (SANB, FNE, BFNE) zeigte die SANB-5 eine klare einfaktorielle Struktur. Dies ist auf den Ausschluss der konstruktnegativ formulierten Items zurückzuführen, die in den anderen Versionen einen zweiten atheoretischen Methodenfaktor bildeten (vgl. Spector et al., 1997). Alle Items wiesen hohe Ladungen auf und können daher als gute Indikatoren für das Konstrukt Bewertungsängstlichkeit angesehen werden. Trotz der hohen Ladungen der Items können diese weder anhand der Inhalte noch psychometrisch als redundant angesehen werden. Es traten zum Beispiel keine nennenswerten Residualkorrelationen auf. Das einfaktorielle Modell der SANB-5 passt sowohl auf die Daten der klinischen als auch auf die Daten der beiden nicht-klinischen Stichproben. Anhand der beiden nicht-klinischen Stichproben konnte sogar skalare Invarianz der SANB-5 gesichert werden. Demnach ist es unerheblich, ob die Items auf Papier oder online im Internet vorgegeben werden. Sie bilden dasselbe Konstrukt ab. Dieser Befund ist insofern bemerkenswert, da die Messinvarianz im Bezug auf den Erhebungs-

modus gerade für Konstrukte aus dem Bereich der negativen Affektivität als kritisch gilt (Buchanan, 2002).

Die Korrelationen mit den Validierungsmaßen liefern Hinweise zur konvergenten und diskriminanten Validität der konstruierten Kurzskala. Die SANB-5 wies substantielle Zusammenhänge zu Selbstbeurteilungsinstrumenten auf, die konzeptuell verwandte Konstrukte erfassen und zeigt geringe oder Nullkorrelationen zu nicht verwandten Konstrukten. Die Korrelationen zwischen SANB-5 und BFI erlauben eine klare Einordnung in das Fünffaktorenmodell der Persönlichkeit (Costa & McCrae, 1992; Goldberg, 1993). Die SANB-5 korreliert hauptsächlich mit *BFI-Neurotizismus*. Auch von anderen Autoren wurde ein hoher Zusammenhang zwischen einer SANB-Version (BFNE) und Neurotizismus gefunden (z. B. Boelen & Reijntjes, 2009). Diese Zuordnung ist plausibel, da Bewertungsängstlichkeit negative Affekte thematisiert und als ein Teilaspekt der Neurotizismus-Facetten Ängstlichkeit angesehen werden kann. Die moderate negative Korrelation mit *BFI-Extraversion* ist ebenfalls konzeptuell nicht unplausibel, da beiden Merkmalen der interpersonelle Kontext gemein ist, der bestimmte Verhaltensweisen auslöst. Bewertungsängstliche könnten soziale Situationen, die von Extravertierten explizit aufgesucht werden, vermeiden, da diese Situationen das Potential für negative Bewertung in sich bergen. Mit den übrigen Faktoren des Fünffaktorenmodells zeigte die SANB-5 keine substantiellen Korrelationen (diskriminante Validität). Mit der Traitskala des STAI, die laut Endler (1975) insbesondere für Angstreaktionen in sozialen Situationen prädiktiv sein soll, zeigte die SANB-5 eine hohe Korrelation (vgl. Vormbrock & Neuser, 1983). Weiterhin wurden differentielle Zusammenhänge der SANB-5 mit den Facetten der Angstsensitivität gefunden. In allen drei Stichproben wurden erwartungsgemäß hohe Korrelationen zur AS-Facetten *Bedenken sozial* gefunden (vgl. Carleton et al., 2007), welche die Furcht vor aversiven sozialen Konsequenzen öffentlich sichtbarer Erregungssymptome erfasst. Personen mit hoher im Vergleich zu Personen mit niedriger Bewertungsängstlichkeit berichten in Bewertungssituationen häufiger Symptome wie Zittern oder Atemlosigkeit (Chen & Drummond, 2008), die von Interaktionspartnern bemerkt und negativ bewertet werden könnten. Die Korrelationen der SANB-5 mit den anderen beiden AS-Facetten fielen deutlich niedriger aus.

Diese korrelativen Befunde replizieren Befunde, die mit englischen Versionen der Skala ermittelt wurden und belegen die Validität der in der vorliegenden Untersuchung konstruierten deutschen Kurzversion der SANB. Darüber hinaus konnte gezeigt werden, dass die Beziehungen im nomologischen Netzwerk durch die Kürzungen nicht beeinträchtigt wurden. Die Validitätskoeffizienten von SANB-5 und SANB zeigen nur unwesentliche Unterschiede. Auch die berechneten Partialkorrelationen zwischen dem Summenwert der 15 eliminierten Items und den Validierungsmaßen bzw. die erklärte Varianz für diese Zusammenhänge zeigen, dass die eliminierten Items im Hinblick auf die Prädiktion der Validierungsmaße keine Aspekte erfassen, die nicht auch durch die gekürzte Version erfasst werden können.

Die Messgenauigkeit der SANB-5 kann angesichts der geringen Anzahl von nur fünf Items als hoch angesehen werden ($.84 < \alpha < .94$). Im Vergleich dazu liegt die interne Konsistenz der Facette Ängstlichkeit des NEO-PI-R mit acht Items bei $\alpha = .82$ (Costa & McCrea, 1992). Dies legt nahe, dass die hohe Messgenauigkeit kein Resultat einer selektiven Konstruktionsstrategie ist – inhaltliche und psychometrische Redundanzen wurden durch den Fokus auf die Inhaltsvalidität vermieden – sondern ein Resultat der Homogenität des erfassten Konstrukts. Da Bewertungsängstlichkeit als Teilaspekt der Ängstlichkeit angesehen werden kann, ist das Konstrukt notwendigerweise homogener im Vergleich zum übergeordneten Konstrukt Ängstlichkeit.

Die hier berichteten Befunde sprechen dafür, dass die SANB-5 eine ökonomische, reliable und valide Erfassung des Konstrukts Bewertungsängstlichkeit erlaubt. Im Vergleich zur SANB hat die SANB-5 eine um 75 % reduzierte Itemzahl und demnach eine um ca. 75 % geringere Bearbeitungsdauer. Trotz der starken Kürzung verringerte sich die Messgenauigkeit lediglich um ca. 7 %. Der Zugewinn an Ökonomie wiegt diese geringfügige Reduktion der Messgenauigkeit auf, zumal die Reliabilität auch nach der Itemselektion noch hoch ist. Die einfach minderungskorrigierte Korrelation zwischen SANB-5 und SANB liegt bei $r = .98$. Die Höhe der Validitätskoeffizienten ändert sich, abgesehen von wenigen praktisch meist unbedeutenden Ausnahmen, durch die Kürzung nicht. SANB-5 und SANB bilden das Konstrukt Bewertungsängstlichkeit demnach fast identisch ab. Die SANB-5 ist aber aufgrund ihrer geringeren Itemzahl deutlich ökonomischer.

Trotz zahlreicher Belege zur Güte der konstruierten Kurzsкала, sollten folgende Einschränkungen nicht unerwähnt bleiben. In der durchgeführten Validierung wurden lediglich Fragebogendaten verwendet. Ein Einfluss der Erhebungsmethode auf die korrelativen Befunde kann daher nicht ausgeschlossen werden. Für zukünftige Studien wäre eine Validierung mittels anderer Datenquellen wünschenswert, um die Validität der SANB-5 weiter abzusichern. Beispielsweise könnten Personen auf der Grundlage einer Verhaltensbeobachtung in sozialen Interaktionssituationen in Gruppen niedrig und hoch sozial Ängstlicher eingeteilt werden. Die SANB-5 sollte diese beiden Gruppen klar diskriminieren können. Für die klinisch-psychologische Forschung wäre der Nachweis relevant, dass die SANB-5 Patientengruppen mit verschiedenen Diagnosen differenzieren kann und zwar in dem Ausmaß, in dem Bewertungsangst Teil der Symptomatik ist. In unserer klinischen Stichprobe war ein solcher Vergleich aufgrund der geringen Anzahl von Sozialphobikern nicht möglich.

Für welche Anwendungsbereiche kann die SANB-5 empfohlen bzw. nicht empfohlen werden? Eine Anwendung in der klinischen Einzelfalldiagnostik ist nicht ratsam. Für diesen Bereich stehen Verfahren zur Verfügung, die soziale Ängstlichkeit oder Phobie störungsspezifisch operationalisieren und für die angemessene Vergleichsstichproben zur Einordnung der Merkmalsausprägung einer Person vorliegen. Als Forschungsinstrument kann die SANB-5 hingegen aufgrund ihrer psychometrischen Güte klar empfohlen werden, zum

Beispiel in der persönlichkeitspsychologischen-, sozialpsychologischen- oder klinisch-psychologischen Forschung. Sie ist aufgrund ihrer ökonomischen Messung – die Bearbeitung der SANB-5 dauert selten länger als eine Minute – besonders für Untersuchungen geeignet, die starken zeitlichen oder monetären Restriktionen unterliegen, wie beispielsweise klinisch-psychologische Studien an universitären psychotherapeutischen Ambulanzen (hohes Aufkommen an Qualifizierungsarbeiten), Onlinestudien oder bei ambulanten Erhebungen mit mobilen Endgeräten, zum Beispiel PDAs. Die Bearbeitungsdauer ist in solchen Studien ein kritischer Faktor, der Teilnahmebereitschaft, Abbrecherquote und Datenqualität, z. B. Häufigkeit fehlender Werte, unvalide Testprotokolle etc., negativ beeinflussen kann (für Onlinestudien zeigen dies z. B. Crawford, Couper, & Lamias, 2001).

Die ökonomische Messung von Persönlichkeitsmerkmalen ist aber nicht nur in diesen Studien von Interesse. Ganz allgemein ermöglicht die alternative Erfassung von Konstrukten über Kurzskalen die Erhebung einer größeren Datenmenge bei gleichbleibender Durchführungsdauer. Aus diesem Grund wäre es prinzipiell wünschenswert, wenn der Ökonomie als Gütekriterium in Zukunft mehr Beachtung geschenkt werden würde. Die Erfassung eines Persönlichkeitsmerkmals mit mehr Items als unter sonst gleichen Bedingungen (ähnliche Reliabilität und Konstruktvalidität von Kurz- und Originalversion) nötig, ist für Skalen, die nur in der psychologischen Forschung genutzt werden, nur schwer vertretbar.

Literaturverzeichnis

- Beauducel, A. & Wittmann, W. W. (2005). Simulation studies on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling, 12*(1), 41–57.
- Boelen, P. A. & Reijntjes, A. (2009). Intolerance of uncertainty and social anxiety. *Journal of Anxiety Disorders, 23*(1), 130–135.
- Buchanan, T. (2002). Online assessment: Desirable or dangerous? *Professional Psychology: Research and Practice, 33*(2), 148–154.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Carleton, R. N., Collimore, K. C. & Asmundson, G. J. (2007). Social anxiety and fear of negative evaluation: Construct validity of the BFNE-II. *Journal of Anxiety Disorders, 21*(1), 131–141.
- Carleton, R. N., McCreary, Norton, P. J. & Asmundson, G. J. (2006). Brief fear of negative evaluation scale – Revised. *Depression and Anxiety, 23*(5), 297–303.
- Chen, V. & Drummond, P. D. (2008). Fear of negative evaluation augments negative affect and somatic symptoms in social-evaluative situations. *Cognition & Emotion, 22*(1), 21–43.
- Costa, P. T. & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO Five-Factor inventory (NEO-FFI): Professional Manual*. Odessa: Psychological Assessment Resources.

- Crawford, S. D., Couper, M. P. & Lamias, M. J. (2001). Web Surveys – Perceptions of Burden. *Social Science Computer Review*, 19(2), 146–162.
- Duke, D., Krishnan, M., Faith, M. & Storch, E. A. (2006). The psychometric properties of the Brief Fear of Negative Evaluation Scale. *Journal of Anxiety Disorders*, 20(6), 807–817.
- Endler, N. S. (1975). A person-situation interaction model for anxiety. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety*, (pp. 145–164). Washington: Hemisphere.
- Faulbaum, F., Prüfer, P. & Rexroth, M. (2009). *Was ist eine gute Frage? – Die systematische Evaluation der Fragenqualität*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gilbert, N. & Meyer, C. (2005a). Fear of negative evaluation and eating attitudes: A replication and extension study. *International Journal of Eating Disorders*, 37(4), 360–363.
- Gilbert, N. & Meyer, C. (2005b). Fear of negative evaluation and the development of eating psychopathology: A longitudinal study among nonclinical women. *International Journal of Eating Disorders*, 37(4), 307–312.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(12), 26–34.
- Hautzinger, M. & Bailer, M. (1993). *Allgemeine Depressions Skala*. Weinheim: Beltz.
- Hinrichsen, H., Waller, G. & Dhokia, R. (2007). Core beliefs and social anxiety in the eating disorders. *Eating and Weight Disorders-Studies on Anorexia, Bulimia, and Obesity*, 12(1), E14–E18.
- Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Kemper, C., Ziegler, M. & Taylor, S. (2009). Überprüfung der psychometrischen Qualität der deutschen Version des Angstsensitivitätsindex-3. *Diagnostica*, 55(4), 223–233.
- Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C. D. (1981). *Das State-Trait-Angst-inventar (STAI). Theoretische Grundlagen und Handanweisung*. Weinheim: Beltz.
- Leary, M. R. (1983). A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, 9(3), 371–375.
- Liebert, R. M. & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 20(3), 975–978.
- Meade, A. W., Johnson, E. C. & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 93(3), 568–592.
- Neuser, J. (2003). Skala „Angst vor negativer Bewertung“ (SANB). In J. Hoyer & J. Margraf (Hrsg.), *Grundlagen und Testverfahren: Angstdiagnostik*. (S. 280–283). Berlin: Springer.
- Oaten, M., Williams, K. D., Jones, A. & Zadro, L. (2008). The effects of ostracism on self-regulation in the socially-anxious. *Journal of Social and Clinical Psychology*, 27(5), 471–504.
- Rammstedt, B. (1997). Die deutsche Version des Big Five Inventory (BFI): Übersetzung und Validierung eines Fragebogens zur Erfassung des Fünf-Faktoren-Modells der Persönlichkeit. Unveröffentlichte Diplomarbeit, Universität Bielefeld.
- Reiss, S. (1991). Expectancy theory of fear, anxiety, and panic. *Clinical Psychology Review*, 11(2), 141–153.

- Rodebaugh, T. L., Woods, C. M., Thissen, D. M., Heimberg, R. G., Chambless, D. L. & Rapee, R. M. (2004). More information from fewer questions: The factor structure and item properties of the original and brief Fear of Negative Evaluation Scale. *Psychological Assessment, 16*(2), 169–181.
- Spector, P. E., van Katwyk, P. T., Brannick, M. T. & Chen, P. Y. (1997). When two factors don't reflect two constructs: How item characteristics can produce artifactual factors. *Journal of Management, 23*(5), 659–677.
- Stanton, J. M., Sinar, E. F., Balzer, W. K. & Smith, P. C. (2002). Issues and strategies for reducing the length of self-report scales. *Personnel Psychology, 55*(1), 167–194.
- Taylor, S. (Ed.) (1999). *Anxiety sensitivity – Theory, research, and treatment of the fear of anxiety*. Mahwah: Erlbaum.
- Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–70.
- Vormbrock, F. & Neuser, J. (1983). Konstruktion zweier spezifischer Fragebögen zur Erfassung von Angst in sozialen Situationen (SANB und SVSS). *Diagnostica, 29*(2), 165–182.
- Watson, D. & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*(4), 448–457.
- Weeks, J. W., Heimberg, R. G., Fresco, D. M., Hart, T. A. & Turk, C. L. (2005). Empirical validation and psychometric evaluation of the Brief Fear of Negative Evaluation Scale in patients with Social Anxiety Disorder. *Psychological Assessment, 17*(2), 179–190.
- Winton, E. C., Clark, D. M. & Edelmann, R. J. (1995). Social anxiety, fear of negative evaluation and the detection of negative emotion in others. *Behaviour Research and Therapy, 33*(2), 193–196.
- Zadro, L., Boland, C. & Richardson, R. (2006). How long does it last? The persistence of the effects of ostracism in the socially anxious. *Journal of Experimental Social Psychology, 42*(5), 692–697.

Korrespondenzadresse: Christoph J. Kemper, GESIS – Leibniz-Institut für Sozialwissenschaften, Mannheim, Quadrat B 2.1, Postfach 12 21 55, 68072 Mannheim; E-Mail: christoph.kemper@gesis.org