

# Zweitveröffentlichung



Hock, Michael; Heim-Dreger, Uwe; Eschenbeck, Heike; u. a.

## Die Diagnostik von Angst und Depression mit dem Mehrdimensionalen Angstinventar für Kinder und Jugendliche (MAI-KJ)

Datum der Zweitveröffentlichung: 22.07.2025

Akzeptiertes Manuskript (Postprint), Zeitschriftenartikel

Persistenter Identifikator: urn:nbn:de:bvb:473-irb-91397x

### Erstveröffentlichung

Hock, Michael; Heim-Dreger, Uwe; Eschenbeck, Heike; u. a. (2020): Die Diagnostik von Angst und Depression mit dem Mehrdimensionalen Angstinventar für Kinder und Jugendliche (MAI-KJ), in: Diagnostica, Göttingen ; Bern [u.a.]: Hogrefe, Jg. 66, Nr. 1, S. 2–13, doi: 10.1026/0012-1924/a000234.

### Verlagshinweis

Diese Artikelfassung entspricht nicht vollständig dem in der Zeitschrift "Diagnostica" veröffentlichten Artikel unter <https://doi.org/10.1026/0012-1924/a000234>. Dies ist nicht die Originalversion des Artikels und kann daher nicht zur Zitierung herangezogen werden.

### Rechtehinweis

Dieses Werk ist durch das Urheberrecht und/oder die Angabe einer Lizenz geschützt. Es steht Ihnen frei, dieses Werk auf jede Art und Weise zu nutzen, die durch die für Sie geltende Gesetzgebung zum Urheberrecht und/oder durch die Lizenz erlaubt ist. Für andere Verwendungszwecke müssen Sie die Erlaubnis der Rechteinhaberinnen und Rechteinhaber einholen.

Für dieses Dokument gilt eine Creative-Commons-Lizenz.



Die Lizenzinformationen sind online verfügbar:

<https://creativecommons.org/licenses/by/4.0/legalcode>

# **Die Diagnostik von Angst und Depression mit dem Mehrdimensionalen Angstinventar für Kinder und Jugendliche (MAI-KJ)**

Michael Hock<sup>1</sup>, Uwe Heim-Dreger<sup>2</sup>, Heike Eschenbeck<sup>2</sup> und Carl-Walter Kohlmann<sup>2</sup>

<sup>1</sup>Institut für Psychologie, Otto-Friedrich-Universität Bamberg

<sup>2</sup>Pädagogische Psychologie und Gesundheitspsychologie, Pädagogische Hochschule  
Schwäbisch Gmünd

Institutsanschriften:

Otto-Friedrich-Universität Bamberg  
Institut für Psychologie  
Markusplatz 3  
96047 Bamberg  
E-Mail: [michael.hock@uni-bamberg.de](mailto:michael.hock@uni-bamberg.de)

Pädagogische Hochschule Schwäbisch Gmünd  
Institut für Humanwissenschaften  
Pädagogische Psychologie und Gesundheitspsychologie  
Oberbettringer Str. 200  
73525 Schwäbisch Gmünd  
E-Mail: [heike.eschenbeck@ph-gmuend.de](mailto:heike.eschenbeck@ph-gmuend.de), [uwe.heim-dreger@ph-gmuend.de](mailto:uwe.heim-dreger@ph-gmuend.de), [carl-walter.kohlmann@ph-gmuend.de](mailto:carl-walter.kohlmann@ph-gmuend.de)

Die finale Version des Dokuments findet sich hier:  
<https://econtent.hogrefe.com/doi/10.1026/0012-1924/a000234>

Zitation::

Hock, M., Heim-Dreger, U., Eschenbeck, H., & Kohlmann, C.-W. (2020). Die Diagnostik von Angst und Depression mit dem Mehrdimensionalen Angstinventar für Kinder und Jugendliche (MAI-KJ) [The assessment of anxiety and depression with the Multidimensional Anxiety Inventory for Children and Adolescents (MAICA)]. *Diagnostica*, 66(1), 2–13. <https://doi.org/10.1026/0012-1924/a000234>

Wir danken Ulrich Schroeders für zahlreiche wertvolle Kommentare zu einer früheren Version der vorliegenden Arbeit.

Schlüsselwörter: Angst; Depression; Besorgnis; Freude; Testkonstruktion

Zusammenfassung: Das allgemeine Modul des „Mehrdimensionalen Angstinventars für Kinder und Jugendliche“ (MAI-KJ) dient der Messung von dispositioneller Angst (indiziert durch Aufgeregtheit und Besorgnis) und Depression (indiziert durch hohe Traurigkeit und niedrige Freude). Der Aufbau des Fragebogens wird vorgestellt. Überprüft werden Faktorenstruktur, psychometrische Eigenschaften und Validität des Verfahrens. Die Analysen basieren vorwiegend auf 2 Stichproben von Kindern und Jugendlichen der Klassenstufen 3 bis 10 (Stichprobe 1:  $N = 2\,594$ , Stichprobe 2:  $N = 7\,339$ ). Die angenommene Struktur konnte mittels explorativer und konfirmatorischer Faktorenanalysen bestätigt werden. Struktur und Messeigenschaften des MAI-KJ waren weitgehend invariant gegenüber Geschlecht und Alter der Kinder. Die Reliabilitäten der Skalen fielen zufriedenstellend bis gut aus (Cronbachs  $\alpha$  zwischen .70 und .88). Die Stabilitäten erwiesen sich als hoch (Korrelationen zwischen .61 und .76, Zeitabstand: 2 bis 4 Wochen). Zahlreiche Hinweise auf die Validität konnten durch Korrelationen mit Fragebogenverfahren zur Diagnostik von Angst und Depression gewonnen werden.

Keywords: anxiety; worry; joy; test construction

Abstract: The general module of the “Multidimensional Anxiety Inventory for Children and Adolescents” (MAICA) is designed to measure dispositional anxiety (indexed by emotionality and worry) and depression (indexed by high dysthymia and low joy). This article describes the construction, the factor structure, the psychometric properties, and the validity of the general module of the MAICA. Analyses are based mainly on 2 samples of children and adolescents attending class levels 3 to 10 (sample 1:  $N = 2\,594$ , sample 2:  $N = 7\,339$ ). The expected factor structure could be substantiated with both exploratory and confirmatory factor analyses, respectively. The structure and the measurement properties of the MAICA proved to be largely invariant with respect to gender and age of the children. The reliabilities of all subscales were satisfactory (Cronbach’s  $\alpha$  between .70 and .88). In addition, all subscales were moderately to highly stable over a period of two to four weeks (correlations between .61 and .76). Convergent and discriminant validity could be established with respect to other anxiety and depression scales.

## Einleitung

Angst gehört bei Kindern und Jugendlichen zu den am häufigsten auftretenden psychischen Problemen. Befragungen an Schülern und Eltern zufolge sind mindestens 10 % der Kinder und Jugendlichen in Deutschland von schwerwiegenden Angstproblemen betroffen (Plück, Döpfner & Lehmkuhl, 2000; Ravens-Sieberer, Wille, Bettge & Erhart, 2007; vgl. Fehm & Fydrich, 2011). Für Prävention und Intervention bei solchen Problemen ist die Verfügbarkeit geeigneter diagnostischer Instrumente eine essenzielle Vorbedingung. Bei Angst ist psychologische Diagnostik besonders wichtig, da Angst zu den „stillen“ Problemen gehört, die sich nicht, wie etwa Hyperaktivität oder Impulsivität, deutlich im offenen Verhalten bemerkbar machen. Sich anbahnende Angstprobleme werden von Erziehenden daher häufig erst sehr spät registriert (nämlich dann, wenn ihre Folgeerscheinungen gravierend werden), in ihrer Bedeutung falsch eingeschätzt oder sogar übersehen (Spinath, 2005). Gut validierte Verfahren zur Angstdiagnostik, die auf dem Erleben von Kindern und Jugendlichen aufbauen, sind daher unverzichtbar, um Angstprobleme frühzeitig aufdecken zu können. Auch für die Erforschung der Bedingungen und Folgen von Angst sind solche Instrumente essenziell.

Mit dem *Mehrdimensionalen Angstinventar für Kinder und Jugendliche* (MAI-KJ; Eschenbeck, Heim-Dreger, Hock & Kohlmann, in Vorb.) wird ein dem Stand der Forschung entsprechendes Verfahren zur Messung von Angst vorgelegt, das für den Einsatz bei 9- bis 16-jährigen Kindern und Jugendlichen (Klassenstufe 3 bis 10) konzipiert ist. Es soll ein theoretisch wie methodisch aktuelles Instrument bereitgestellt werden, das zudem vergleichsweise ökonomisch ist, einen breiten Anwendungsbereich aufweist und dessen Reliabilität auch für die Einzelfalldiagnostik ausreichend ist. Im Unterschied zu bereits vorliegenden Verfahren, wie zum Beispiel dem *Angstfragebogen für Schüler* (Wieczerkowski, Nickel, Janowski, Fittkau & Rauer, 1974; Wieczerkowski, Nickel, Janowski, Fittkau, Rauer & Petermann, 2016), der trotz bekannter kritischer Punkte (z. B. undifferenzierte Erfassung von Angst; siehe Rost & Haferkamp, 1979) seit über 40 Jahren unverändert im Einsatz ist, wurde besonderer Wert auf (a) eine hinsichtlich Depression diskriminant valide Erfassung zentraler Komponenten der Angst, (b) eine umfassende Prüfung der faktoriellen Struktur des Verfahrens mittels explorativer und konfirmatorischer Faktorenanalysen, die für Itemantworten angemessen sind,

und (c) den Nachweis der Messinvarianz gegenüber Geschlecht und Alter, gelegt. Letzterer ist speziell bei Verfahren für Kinder und Jugendliche, die eine große Altersspanne abdecken sollen, essenziell. Im intendierten Anwendungsbereich stellen die genannten Eigenschaften Alleinstellungsmerkmale des MAI-KJ dar.

Das MAI-KJ ist in vier Module gegliedert, die unabhängig voneinander bearbeitet werden können, so dass, je nach Untersuchungszweck, ein ökonomisches Screening im Hinblick auf emotionale Auffälligkeiten, z. B. für Untersuchungen im Klassenverband, oder eine intensive Einzelfalldiagnostik, z. B. im Rahmen einer schulpsychologischen Beratung, durchgeführt werden kann. Alle Module sind Selbstberichtsverfahren, mit denen Angst und mit Angst assoziierte Emotionen als längerfristig stabile Dispositionen bestimmt werden sollen. In Modul 1, das als Screening-Instrument konzipiert ist, wird erfasst, wie sich ein Kind im Allgemeinen fühlt. Gemessen werden hier also generelle (situationsübergreifende) Dispositionen. Neben angstbezogenen Gefühlen und Kognitionen (Aufgeregtheit und Besorgnis) werden auch die depressionsbezogenen Emotionen Freude und Traurigkeit diagnostiziert. Damit soll das Modul die Trennung von Angst und Depression ermöglichen, was bislang kein anderes für Kinder und Jugendliche konzipiertes deutschsprachiges Verfahren leistet. Das Modul kann für die Beantwortung theoretischer wie auch praktisch-angewandter Fragen eingesetzt werden, bei denen es um die Ausprägung oder Veränderung der genannten Dispositionen geht. Während das Einsatzgebiet des ersten Moduls sehr allgemein ist, sind die verbleibenden Module 2 bis 4 primär für Fragen der Pädagogischen Psychologie, insbesondere der Schulpsychologie, entworfen. Diese Module sind als Situations-Reaktionsinventare aufgebaut. Hier wird jeweils eine konkrete schulische Situation geschildert, in die sich die Kinder hineinversetzen sollen. In den Situationen werden leistungs- bzw. bewertungsbezogene und soziale Merkmale variiert: eine Klassenarbeit schreiben (Modul 2), in der Pause mit anderen Kindern zusammen sein (Modul 3), vor der Klasse abgefragt werden (Modul 4). Dies ermöglicht es, neben dem allgemeinen Befinden auch situationspezifische Profilierungen angstassoziierter Zustände, etwa affektive Blockaden bei „Auftritten“ oder die Angst, sich vor Anderen zu blamieren, zu bestimmen. Solche detaillierten Informationen sind vor allem in Beratungskontexten nützlich.

Die vorliegende Arbeit konzentriert sich auf das Modul 1, also den *allgemeinen Teil* des MAI-KJ.

### **Erfassung von Angst und Depression**

Wie erwähnt, werden im allgemeinen Teil des MAI-KJ neben Angst, bei der zwischen Aufgeregtheit und Besorgnis unterschieden wird, auch Freude und Traurigkeit gemessen. Die Einbeziehung von Freude und Traurigkeit verdankt sich dem Versuch, nach dem Vorbild des *State-Trait-Angst-Depressions-Inventars* (STADI; Laux, Hock, Bergner-Köther, Hodapp & Renner, 2013), das für Erwachsene konzipiert ist, zwischen Angst (indiziert durch Aufgeregtheit und Besorgnis) und Depression (indiziert durch hohe Traurigkeit und niedrige Freude) zu differenzieren. Eines der Hauptanliegen des STADI, das im allgemeinen Teil des MAI-KJ übernommen wurde, ist es, die beiden engen Verwandten Angst und Depression diskriminativ valide zu erfassen. Tatsächlich werden in vielen Fragebogen zur Erfassung von Angst bzw. Depression Items aus beiden Bereichen miteinander vermengt, was beim MAI-KJ vermieden werden sollte.

Gemeinsamkeiten von und Unterschiede zwischen Angst und Depression sind seit langem ein wichtiges Thema der Persönlichkeitsdiagnostik und der klinisch-psychologischen Diagnostik (Watson, 2009). Hierbei geht es nicht nur um die Abgrenzung verschiedener klinischer Störungsbilder, sondern auch um interindividuelle Differenzen in angst- und depressionsbezogenen Zuständen und Eigenschaften in nichtklinischen Populationen. Spezifika von Angst und Depression sind dabei im Hinblick auf die Grundlagen und Manifestationen von emotionalen Beeinträchtigungen und Störungen, deren Folgen wie auch hinsichtlich jeweils angemessener Interventionen bedeutsam. Im Hinblick auf präventive Maßnahmen bei Kindern und Jugendlichen, bei denen es darauf ankommt, Vorläufer massiver emotionaler Beeinträchtigungen frühzeitig zu erkennen, ist die Diagnostik interindividueller Differenzen in einem breiten Bereich ein essenzieller Forschungsgegenstand (z. B. Cummings, Caporino & Kendall, 2014). Um Angst und Depression bei Kindern und Jugendlichen möglichst gut separieren zu können, werden im MAI-KJ Ideen aus dem Tripartite-Modell (Clark & Watson, 1991) aufgegriffen.

Ausgangspunkt dieses Modells ist die sehr hohe Korrelation von Angst- und Depressionsmaßen, die in einer Vielzahl von Untersuchungen belegt wurde, und sich im klinisch-psychologischen Bereich in einer deutlichen Komorbidität beider Störungen wiederfindet (für den Kinder- / Jugendbereich siehe Cummings et al., 2014; Zavaglia & Bergeron, 2017; vgl. auch Spence, 2017). Die Korrelation ist auch für Instrumente, die inhaltstvalid zusammengestellt wurden (bei denen die Items sich also eindeutig dem Angst- oder dem Depressionsbereich zuordnen lassen), häufig so hoch, dass die Trennbarkeit von Angst und Depression mit Hilfe von Selbstberichten bezweifelt wurde (zusammenfassend Renner, Hock, Bergner-Köther & Laux, 2018). Im Tripartite-Modell wird die starke Assoziation von Angst- und Depressionsmaßen darauf zurückgeführt, dass beide einen *unspezifischen* Einflussfaktor teilen, der als negative Affektivität (NA) bezeichnet wird (Watson, Clark & Tellegen, 1988). NA repräsentiert einen sehr breit definierten Persönlichkeitsfaktor, der die Tendenz einer Person beschreibt, aversive emotionale Zustände (Ärger, Furcht, Trauer, Ekel, Verachtung usw.) zu manifestieren. NA entspricht damit im Wesentlichen dem Neurotizismusfaktor aus dem „Big Five“-Ansatz. Darüber hinaus werden zwei *spezifische* Merkmale postuliert, die relativ unabhängig voneinander variieren und Angst und Depression voneinander abheben. Das spezifische Merkmal der Angst wird in starker körperlicher Erregung (*physiological hyperarousal*) gesehen. Das spezifische Merkmal der Depression ist dagegen Anhedonie, also der Mangel an positivem Affekt. Für eine distinkte Erfassung von Angst und Depression durch Fragebogen werden damit Itemgruppen benötigt, die, neben NA, körperliche Erregung und positiven Affekt ansprechen.

Das Tripartite-Modell repräsentiert unzweifelhaft einen, auch aufgrund seiner Einfachheit, konzeptuell überzeugenden und darüber hinaus empirisch gut belegten Ansatz zur Trennung von Angst und Depression (Mineka, Watson & Clark, 1998; Watson, 2009). Das Modell wurde auch im Bereich der Forschung zu Angst und Depression bei Kindern und Jugendlichen aufgegriffen (Anderson & Hope, 2008; Lonigan, Philips & Hooe, 2003).

Allerdings kann unserer Auffassung nach, die wiederum dem STADI-Modell folgt (Renner et al., 2018; vgl. auch Brodbeck, Abbott, Goodyer & Croudace, 2011), der unspezifische NA-Faktor, welcher für die starke Kovariation von Angst und Depression

verantwortlich sein soll, durchaus spezifischer gefasst werden. An die Stelle von NA werden mit Besorgnis und Traurigkeit zwei spezifischere Konstrukte gesetzt, von denen das erste dem Angst- und das zweite dem Depressionsbereich zugeordnet wird.

In der allgemein-psychologischen und persönlichkeitspsychologischen Angstforschung hat sich seit den Arbeiten von Liebert und Morris (1967) die Differenzierung zwischen einer kognitiven und einer – im engeren Sinne – emotionalen Angstkomponente durchgesetzt (siehe z. B. Krohne, 2010). Diese Differenzierung betraf ursprünglich die Prüfungsangst, ist aber prinzipiell für alle Bereiche, in denen sich Angst manifestieren kann, bedeutsam. Die kognitive Komponente wird als Besorgnis bezeichnet und ist durch Gedanken charakterisiert, die um mögliche aversive Ausgänge bedrohlicher Situationen und deren weitere Konsequenzen für die Person kreisen. Die emotionale Komponente wird als Aufgeregtheit bezeichnet und meint die Wahrnehmung körperlicher Erregung. Aufgeregtheit umfasst relativ unspezifische Angstzeichen, wie etwa Anspannung, Nervosität oder Unruhe, aber auch umgrenzte körperliche Symptome wie z. B. Zittern oder Schwitzen. Aufgeregtheit ist damit identisch mit der im Tripartite-Modell postulierten spezifischen Angstkomponente. Besorgnis und körperliche Erregungssymptome (Aufgeregtheit) werden auch in der klinisch-psychologischen Literatur als zentrale Bestandteile der Angst thematisiert. So gehören sie sowohl im Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013) wie in der International Statistical Classification of Diseases and Related Health Problems (ICD-10; WHO, 1993) zu den Kernkriterien der generalisierten Angststörung.<sup>1</sup>

Besorgnis und Aufgeregtheit lassen sich nicht nur faktorenanalytisch differenzieren, sie zeigen einen jeweils spezifischen zeitlichen Verlauf, sind unterschiedlich sensitiv für Rückmeldungen und differenziell mit den Leistungen bei komplexen Aufgaben assoziiert

---

<sup>1</sup> Speziell in der klinisch-psychologischen Forschung zu Angst und Depression ist mit Rumination ein weiteres Konzept prominent, das dem der Besorgnis insofern ähnelt, als hier ebenfalls Kognitionen mit negativem Inhalt angesprochen werden. Während sich Besorgnis jedoch auf antizipierte Bedrohungen bezieht, welche die Personen zu kontrollieren versucht, kreisen Gedanken bei der Rumination um vergangene oder gegenwärtige die Person betreffende aversive Ereignisse, die z. B. Verlust- oder Versagenserlebnisse beinhalten und deren Sinn die Person tiefer verstehen möchte (Nolen-Hoeksema, Wisco & Lyubomirsky, 2008). Der diagnostische Status der Rumination ist noch offen. Im Unterschied zur Besorgnis wird sie eher der Depression zugeordnet als der Angst oder es wird angenommen, dass sie für Angst und Depression gleichermaßen charakteristisch ist. Eine ausführliche Diskussion der Unterschiede zwischen Besorgnis und Rumination liefern Nolen-Hoeksema et al. (2008). Um beim MAI-KJ eine eindeutige Gliederung in einen Angst- und einen Depressionsbereich zu gewährleisten, fokussiert das Verfahren die Besorgnis, schließt Ruminationsitems also aus.

(zusammenfassend Zeidner, 1998). Es ist daher sinnvoll, beide Komponenten der Angst separat zu erfassen. Die Unterscheidung kann darüber hinaus auch bei Interventionen hilfreich sein, da die Maßnahmenswerpunkte bei einem Angstproblem (z. B. Prüfungsangst) darauf abgestimmt werden sollten, ob bei einer Person Besorgnis oder Aufgeregtheit im Vordergrund steht.

Auch für die Depression erscheint es sinnvoll, den NA-Anteil einzugrenzen, hier auf Gefühle der Traurigkeit und Niedergeschlagenheit (Renner et al., 2018). Sicherlich sind nicht alle Emotionen, die unter NA fallen, für Depression gleichermaßen charakteristisch. Während eine traurige, gedrückte bzw. niedergeschlagene Stimmung zu den affektiven Kernmerkmalen der Depression gehört, die in den meisten Definitionsvorschlägen auftaucht, werden andere negative Emotionen wie z. B. Ärger oder Scham als begrifflich von der Depression unabhängige und allenfalls mit ihr assoziierte Zustände betrachtet.

Mit dem allgemeinen Teil des MAI-KJ sollen damit vier Konstrukte erfasst werden: Aufgeregtheit, Besorgnis, Freude und Traurigkeit. Während Aufgeregtheit und Besorgnis als zentrale Facetten der Angst betrachtet werden, werden Traurigkeit und der Mangel an Freude (Anhedonie) der Depression zugeordnet.

### **Ziele und Fragestellungen**

Mit der Konstruktion des MAI-KJ wurde das Ziel verfolgt, Skalen zu erstellen, die es ermöglichen, die genannten Konstrukte bei Kindern und Jugendlichen reliabel und valide zu erfassen. Die Skalen sollten längerfristig stabile Dispositionen messen, was im intendierten Anwendungsbereich wichtiger ist als die Messung momentaner emotionaler Zustände. Wie bei ähnlich konzipierten Instrumenten war es nicht das Ziel, klinisch-psychologische Diagnosen abzubilden, vielmehr sollten Komponenten der Angst und Depression bestimmt werden, deren Variation sowohl im nichtklinischen wie im klinischen Bereich relevant ist.

Was inhaltliche Gesichtspunkte betrifft, wurde darauf geachtet, Items zu formulieren, die sich semantisch eindeutig einem der vier anvisierten Bereiche zuordnen lassen. Diese Zuordnung wurde zunächst mit explorativen Faktorenanalysen geprüft. Die Ergebnisse dieser Analysen führten zu diversen Modifikationen des MAI-KJ, dessen Endversion sodann mit konfirmatorischen Modellen untersucht wurde.

Die zentrale Annahme besteht darin, dass – entsprechend den Untersuchungen zum STADI mit Erwachsenen (Laux et al., 2013; Renner et al., 2018) – Aufgeregtheit, Besorgnis, Freude und Traurigkeit auch bei Kindern und Jugendlichen distinkte Konstrukte repräsentieren. Diese Annahme wurde mit Hilfe eines konfirmatorischen Modells getestet, in dem jedes interessierende Konstrukt durch *einen* Faktor repräsentiert wird, der einer Untermenge der Items gemeinsam ist (siehe Abbildung 1, linker Teil). Aus konzeptuellen Gründen ist es naheliegend, darüber hinaus auch ein verwandtes Modell zu betrachten, in dem zusätzlich Angst und Depression als korrelierte Faktoren zweiter Ordnung eingeführt werden (siehe Abbildung 1, rechter Teil). Für dieses Angst-Depressionsmodell werden die Freude-Items invertiert, so dass der entsprechende Faktor Anhedonie reflektiert. Eine gute Passung dieses Modells würde eine Grundlage für die Verwendung höher aggregierter Angst- und Depressionsskalen liefern, die in Anwendungskontexten nützlich sein können.

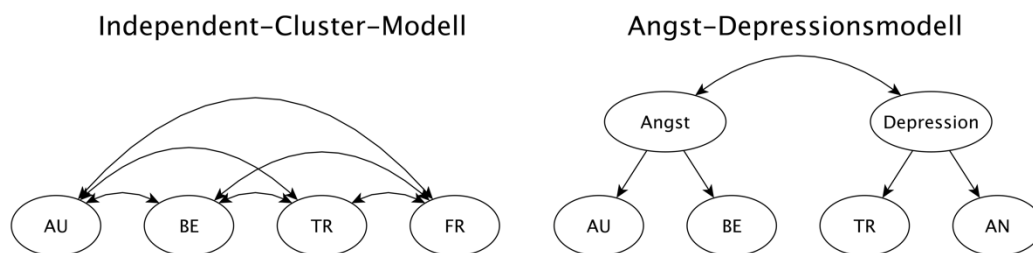


Abbildung 1. Illustration der konfirmatorischen Faktorenmodelle. AU = Aufgeregtheit, BE = Besorgnis, TR = Traurigkeit, FR = Freude, AN = Anhedonie. Die Faktoren werden durch 6 (AU), 7 (BE), 5 (TR) bzw. 7 (FR/AN) Items indiziert, die der Übersichtlichkeit halber nicht dargestellt sind. Für die Gewinnung von AN-Indikatoren wurden die FR-Items invertiert.

Für ein Verfahren, das wie das MAI-KJ einen großen Altersbereich abdecken soll, stellt sich die Frage, ob in verschiedenen Altersgruppen jeweils die gleichen Merkmale gemessen werden. Besteht für das MAI-KJ Messinvarianz bei Kindern unterschiedlichen Alters? Da für Angst und Depression in der Literatur (z. B. Costa, Terracciano & McCrae, 2001) recht konsistente Geschlechterunterschiede dokumentiert sind (Frauen manifestieren höhere Werte als Männer), deren Grundlage noch unklar ist (handelt es sich um genuine Differenzen in den Merkmalsausprägungen, um Unterschiede in der Interpretation von Items oder vielleicht nur

um reine Antworttendenzen?), kombinierten wir die Prüfung der altersbezogenen mit der Prüfung der geschlechterbezogenen Messinvarianz.

Im Idealfall messen die Items eines Fragebogens die anvisierten Konstrukte in allen Gruppen in der gleichen Weise und mit gleicher Genauigkeit, was impliziert, dass Faktorladungen, Schwellenparameter und Residualvarianzen der Items nicht von der Gruppenzugehörigkeit abhängen. Prüfen lässt sich Messinvarianz, indem Mehrgruppen-Faktorenmodelle berechnet werden, in denen die genannten Parameter über die Gruppen als gleich spezifiziert werden. Da solche Idealforderungen in der Praxis allerdings selten vollständig erfüllt sind, werden zur Prüfung der Messinvarianz im Allgemeinen eine Reihe weniger restriktiver Mehrgruppen-Faktorenmodelle betrachtet und im Hinblick auf ihre Passung miteinander kontrastiert. Durch den Vergleich dieser Modelle ist es dann möglich, Quellen der (Non-) Invarianz genauer zu lokalisieren.

Die vorliegende Untersuchung zielte primär auf die Prüfung der Faktorenstruktur und Messinvarianz des allgemeinen Teils des MAI-KJ ab. Darüber hinaus sollten auch Hinweise auf das Vorliegen konvergenter und diskriminanter Validität der MAI-KJ-Skalen gewonnen werden. Hierfür wurden Korrelationen zwischen diesen Skalen und Skalen aus vorliegenden Verfahren zur Messung von Angst und Depression betrachtet. Die Verfahren und Erwartungen sind (zusammen mit Stichprobenbeschreibungen) im elektronischen Supplement (ESM) 1 zusammengefasst.

## **Methode**

### **Stichproben und Durchführung**

Die Faktorenanalysen basieren auf Daten aus zwei Stichproben von Kindern und Jugendlichen der Klassenstufen 3 bis 10 (Alter: 8 bis 19 Jahre; Mädchen / Jungen jeweils 50 %), die im Jahr 2012 (Stichprobe 1,  $N = 2\,660$ ) und in den Jahren 2013 und 2014 (Stichprobe 2,  $N = 7\,620$ ) hauptsächlich in Süddeutschland rekrutiert worden waren. Fragebogen, bei denen mehr als 5 % der Items unbeantwortet geblieben waren, wurden aus den Analysen ausgeschlossen. Dies betraf häufig Kinder, die versehentlich eine ganze Seite überblättert hatten. Durch den Fallausschluss reduzierten sich die Stichprobenumfänge auf  $N = 2\,594$  (Alter:  $M = 11.6$ ,  $SD = 2.1$ ; 50 % Mädchen) bzw.  $N = 7\,339$  (Alter:  $M = 11.5$ ,  $SD = 2.3$ ; 50 %

Mädchen). Die Untersuchungen fanden generell im Klassenverband statt und nahmen ein bis zwei Schulstunden in Anspruch. Für die Bestimmung von Zusammenhängen zwischen den MAI-KJ-Skalen und anderen Instrumenten bearbeiteten die Schülerinnen und Schüler neben dem MAI-KJ zusätzlich mindestens ein weiteres Verfahren.

Die Stichproben und das Vorgehen bei der Behandlung fehlender Werte sind im ESM 2 genauer beschrieben.

### **Items und Itemselektion**

Beim MAI-KJ werden die befragten Kinder und Jugendlichen instruiert, ihre Gefühle und Gedanken zu beschreiben und auf den vorgegebenen Antwortoptionen die am besten zutreffende anzukreuzen. Die Items des MAI-KJ sind jeweils mit einer vierstufigen Häufigkeitsskala versehen. Die Antwortoptionen sind 1 = (*fast*) *nie*, 2 = *manchmal*, 3 = *oft* und 4 = (*fast*) *immer*. Voruntersuchungen hatten gezeigt, dass auch jüngere Kinder mit diesem Antwortformat gut zurechtkommen (siehe ESM 2).

Die Items wurden sukzessiv aus einer umfassenden Menge selektiert. Die erste Version des Verfahrens enthielt insgesamt 187 Items, wovon 55 auf den allgemeinen Teil (Modul 1) entfielen, die letzte Version enthielt 88 Items, davon 25 im allgemeinen Teil. Neben inhaltlichen Gesichtspunkten wurden für die Auswahl der Items deren Verteilungscharakteristiken, Faktorladungen und faktorielle Komplexität in explorativen Analysen berücksichtigt. Bei der Auswahl der Items und deren Zuordnung zu Skalen orientierten wir uns an den folgenden Kriterien: (a) die höchste Ladung eines Items auf einem der Faktoren (Hauptladung) ist größer als .40; (b) die faktorielle Komplexität des Items (Hofmann, 1978) ist gering (kleiner als 2; dies impliziert, dass die Nebenladungen deutlich kleiner ausfallen als die Hauptladungen); (c) die Itemcharakteristiken sind monoton steigend; (d) die Reliabilität der Skala erreicht einen Wert von mindestens .80. Zusätzlich wurde darauf geachtet, dass die Itemformulierungen auch für jüngere Kinder (dritte Klasse) verständlich sind. Items, bei denen gehäuft Nachfragen der Kinder zu beobachten waren, wurden eliminiert. Die Itemstämme der Endversion des MAI-KJ sind Tabelle 1 zu entnehmen.

## Faktorenanalysen

Explorative Faktorenanalysen wurden nach dem Common-Factors-Modell berechnet. Hierzu wurden die Daten aus Stichprobe 1 herangezogen. Um die Zusammenhänge der den Itemantworten zugrunde liegenden quantitativen Variablen möglichst genau zu rekonstruieren und das Auftreten sogenannter Schwierigkeitsfaktoren (artifizieller Faktoren, deren Existenz sich allein Schwierigkeitsdifferenzen zwischen den Items verdankt) zu vermeiden, wurden die Analysen auf der Basis polychorischer Korrelationen durchgeführt.

Konfirmatorische Faktorenanalysen wurden mit den Daten der Stichprobe 2 durchgeführt. Für die Berechnungen mit dem R-Paket *Lavaan* (Rosseel, 2012) verwendeten wir robuste Weighted Least Squares-Schätzer (WLSMV), was im Wesentlichen einer Analyse polychorischer Korrelationen entspricht. Wie inzwischen üblich, orientierten wir uns bei der Beurteilung der Modellpassung an approximativen (deskriptiven) Fitindices (Beauducel & Wittmann, 2005; Hu & Bentler, 1999), da  $\chi^2$ -Tests bei großen Stichproben zu sensitiv auf geringfügige Abweichungen zwischen Modell und Daten reagieren. Berechnet wurden ein inkrementeller (CFI; Werte größer als .95 indizieren gute Passung) und ein absoluter Fitindex (RMSEA; Werte kleiner als .06 indizieren gute Passung).

Wie erwähnt, wurden zwei konfirmatorische Modelle berechnet (siehe Abbildung 1). Dem ersten konfirmatorischen Modell, das im linken Teil der Abbildung 1 dargestellt ist, wurde eine sogenannte Independent-Cluster- (IC-) Struktur zugrunde gelegt (McDonald, 1999). In einem IC-Modell wird jedes interessierende Konstrukt durch *einen* Faktor repräsentiert, der einer Untermenge der Items gemeinsam ist. Zudem werden die Antworten auf jedes Item nur durch jeweils *einen* Faktor bestimmt. Das zweite Merkmal unterscheidet ein IC-Modell von einem explorativen Modell, in dem jedes Item von allen Faktoren abhängt. Die in einem IC-Modell angenommene Abhängigkeit jedes Items von genau einem Faktor impliziert, dass die Kovariationen unter den Items vollständig durch ihre Abhängigkeit von diesem Faktor erklärt werden können. Demgegenüber hängen die Itemkovariationen in einem explorativen Modell von allen Faktoren ab. Annahmen bezüglich der Existenz separierbarer („independent“) Gruppen („cluster“) von Items sind damit in einem IC-Modell, das wesentliche Elemente von Thurstones Forderung nach Einfachstruktur operationalisiert (siehe McDonald, 1999, Kapitel

9), deutlich restriktiver als in einem explorativen Modell und stellen recht hohe Anforderungen an die faktorielle Reinheit der Items.

Das darüber hinaus geprüfte Angst-Depressionsmodell (Abbildung 1, rechter Teil), in dem, zusätzlich zu den vier Faktoren erster Ordnung, Angst und Depression als (korrelierte) Faktoren zweiter Ordnung betrachtet werden, macht etwas strengere Annahmen im Hinblick auf die Kovariationen der Faktoren erster Ordnung als das IC-Modell. Während diese im IC-Modell völlig frei variieren können, schränkt das Angst-Depressionsmodell die Kovarianzen zwischen den Faktoren erster Ordnung ein, da z. B. Aufgeregtheit und Besorgnis jeweils über den gleichen Pfad (nämlich den ungerichteten Pfad zwischen Angst und Depression) mit Traurigkeit (oder Freude bzw. Anhedonie) verknüpft sind.

### **Prüfung der Messinvarianz**

Die Messinvarianz wurde separat für das Geschlecht und das Alter der Kinder geprüft. Für die Bewertung der altersbezogenen Messinvarianz teilten wir die Stichprobe 2 anhand der Klassenstufe in drei Gruppen ein. Die Einteilung wurde so vorgenommen, dass hinreichend große und möglichst ähnliche Gruppenumfänge resultierten. Außerdem erschien es sinnvoll, Grundschul Kinder (also die Klassenstufen 3 und 4) einer eigenen Gruppe zuzuordnen. Zusammengefasst wurden daher Kinder der Klassenstufen 3 und 4, 5 bis 7 sowie 8 bis 10, was zu Gruppenumfängen von  $n = 2\,538$ ,  $2\,834$  und  $1\,967$  führte. Wir verwendeten die Klassenstufe als Stellvertreter für das Alter, da die Klassenstufe (die mit dem Alter nahezu perfekt korreliert war,  $r = .96$ ) den sozialen und emotionalen Entwicklungsstand von Kindern vermutlich besser repräsentiert als das chronologische Alter.

Eine vorab vorgenommene Inspektion der für die Gruppen separat berechneten Vier-Faktorenmodelle, für die wiederum die WLSMV-Schätzung verwendet wurde, zeigte keine Auffälligkeiten: Es ergab sich jeweils eine gute Passung der Modelle mit substantziellen Ladungen aller Indikatorvariablen. Für die formelle Prüfung der Messinvarianz wurde die Theta-Parametrisierung verwendet, in der im Unterschied zur Delta-Parametrisierung die Residualvarianzen (außer in einer Gruppe) als freie Parameter behandelt werden können

(Millsap & Yun-Tein, 2004; Muthén & Asparouhov, 2002).<sup>2</sup> Die Passung der diversen Mehrgruppenmodelle (konfigurale, schwache, starke, strikte Invarianz), die im ESM 4 näher beschrieben werden, wurde wiederum mit Hilfe des CFI und des RMSEA beurteilt, da Likelihood-Ratio-Tests bei großen Stichproben auch bei Modellvergleichen zu sensitiv reagieren. Der CFI wurde in mehreren Simulationsstudien übereinstimmend als vergleichsweise guter Kennwert für diesen Einsatzzweck ausgezeichnet (Chen, 2007; Cheung & Rensvold, 2002; Meade, Johnson & Braddy, 2008). Neben der Betrachtung der absoluten Höhe diverser approximativer Fitindices, wurde auch – in Analogie zu Likelihood-Ratio-Tests – die Inspektion von Veränderungen dieser Indices in zunehmend restriktiveren Modellen vorgeschlagen. Als Orientierungsmarke für bedeutsame Verschlechterungen eines restriktiven (z. B. schwache Invarianz) gegenüber einem weniger restriktiven Modell (z. B. konfigurale Invarianz) schlagen Cheung und Rensvold (2002) für den CFI eine Abnahme und für den RMSEA eine Erhöhung von .01 oder mehr vor. Meade et al. (2008) kommen zu etwas strikteren Grenzwerten ( $\Delta\text{CFI} = .002$ ,  $\Delta\text{RMSEA} = .007$ ). Die genannten Studien konzentrieren sich allerdings auf Modelle für kontinuierliche Variablen. Für die Prüfung der Messinvarianz bei Faktorenmodellen für kategoriale Daten (z. B. WLSMV-Schätzung) mit Hilfe approximativer Fitindices liegen zurzeit nur wenige (eher kritische) Erfahrungen vor (siehe hierzu Sass, Schmitt & Marsh, 2014), weshalb wir uns im Folgenden vornehmlich auf deren absolute Höhe stützen.

## **Ergebnisse**

### **Explorative Faktorenanalyse (Stichprobe 1)**

Die Ergebnisse zeigten eine klar interpretierbare Vier-Faktorenstruktur, die Items wurden sehr genau den postulierten Aufregungs-, Besorgnis-, Traurigkeits- und Freudefaktoren

---

<sup>2</sup> Bei (Mehrgruppen-) Faktorenanalysen für ordinale Variablen stellt sich die Frage, ob die sog. Delta- oder die Theta-Parametrisierung verwendet werden soll. Beide Arten der Parametrisierung beinhalten Restriktionen, die, zusammen mit weiteren Annahmen (siehe Millsap & Yun-Tein, 2004) die Identifizierbarkeit der Modelle gewährleisten sollen, und betreffen die den manifesten kategorialen Itemantworten zugrunde liegenden latenten kontinuierlichen Antwortvariablen. Der Unterschied zwischen beiden besteht darin, dass bei der Delta-Parametrisierung die Varianzen, bei der Theta-Parametrisierung dagegen die Residualvarianzen der kontinuierlichen Antwortvariablen als Parameter fungieren. Die Parameter (Varianzen oder Residualvarianzen) werden in einer Gruppe auf 1 fixiert und können dann in den anderen Gruppen geschätzt werden. Da bei Mehrgruppenanalysen u.a. die Gleichheit der Residualvarianzen interessiert und geprüft werden soll, ist die Theta-Parametrisierung der Delta-Parametrisierung hier vorzuziehen.

zugeordnet. Vorgehensweise und Resultate sind detailliert im ESM 3 beschrieben. Die Lösung lieferte die Basis für die Itemzuordnung in der konfirmatorischen Faktorenanalyse.

### **Konfirmatorische Faktorenanalysen (Stichprobe 2)**

Der  $\chi^2$ -Wert für das Vier-Faktorenmodell (IC-Modell, siehe Abbildung 1) fällt zwar signifikant aus ( $\chi^2 = 4\,313.3$ ,  $df = 269$ ), Fitindices verwiesen aber auf eine insgesamt gute Passung: CFI = .965, RMSEA = .045, 90-Prozent-Konfidenzintervall für den RMSEA zwischen .044 und .046. Die Lösung ist in Tabelle 1 zusammengestellt. Wie zu erkennen ist, waren alle standardisierten Ladungen relativ hoch und fielen nicht unter einen Wert von .46; die meisten Ladungen lagen über einem Wert von .60. Die Korrelationen zwischen den Faktoren waren für Skalen, die sich auf negative Affektivität beziehen, recht deutlich ( $r$ s zwischen .72 und .77). Freude war stärker mit Traurigkeit ( $r = -.60$ ) als mit Aufgeregtheit ( $r = -.35$ ) und Besorgnis ( $r = -.36$ ) assoziiert.

Das Korrelationsmuster der Faktoren lässt ein Angst-Depressionsmodell recht plausibel erscheinen. Wie erwähnt, werden in diesem Modell zusätzlich zum IC-Modell zwei Faktoren zweiter Ordnung angenommen, nämlich ein Angstfaktor, der Aufgeregtheit und Besorgnis beeinflusst, sowie ein Depressionsfaktor, der Traurigkeit und Freude (bzw. Anhedonie) beeinflusst (Abbildung 1). Auch dieses Modell wurde geprüft, wobei die Ladungen der jeweils zwei abhängigen Faktoren erster Ordnung (Besorgnis und Aufgeregtheit; Traurigkeit und Anhedonie) zur Identifikation gleichgesetzt wurden. Die Passung des Angst-Depressionsmodells war schlechter als die des IC-Modells, insgesamt jedoch noch als befriedigend anzusehen,  $\chi^2 = 6\,361.0$ ,  $df = 272$ , CFI = .947, RMSEA = .055, 90-Prozent-Konfidenzintervall für den RMSEA zwischen .054 und .056. Die Korrelation der beiden Faktoren zweiter Ordnung betrug .77.

*Tabelle 1.* Standardisierte Ladungen der Items und Korrelationen der Faktoren in der konfirmatorischen Faktorenanalyse des MAI-KJ (allgemeiner Teil)

	AU		BE		TR		FR
	Item-Nr. und Faktorladung						
2	.52	3	.65	4	.67	1	.74
5	.53	7	.73	8	.74	6	.79
12	.71	10	.53	14	.76	11	.71
16	.54	13	.46	18	.71	15	.75
20	.64	17	.74	26	.79	19	.86
22	.57	21	.64			23	.86
		25	.65			27	.80
	Faktor- und Skalengkorrlationen						
AU			.77		.72		-.35
BE	.57				.72		-.36
TR	.52		.55				-.60
FR	-.28		-.29		-.49		

*Anmerkungen:* AU = Aufgeregtheit, BE = Besorgnis, TR = Traurigkeit, FR = Freude. Untere Teiltabelle: Faktorkorrelationen sind über der Diagonalen, Skalengkorrlationen unter der Diagonalen angegeben.  $N = 7\ 339$ .

Items (sortiert nach Faktor) AU: 2 = „Ich fühle mich angespannt.“ 5 = „Mein Herz klopft stark.“ 12 = „Ich habe ein unangenehmes Gefühl im Bauch.“ 16 = „Ich bin unruhig.“ 20 = „Ich bin nervös.“ 22 = „Meine Hände oder Beine zittern vor Aufregung.“

BE: 3 = „Ich habe Angst, etwas falsch zu machen.“ 7 = „Ich mache mir Sorgen, mich zu blamieren.“ 10 = „Wenn mir etwas nicht gelungen ist, muss ich immer wieder daran denken.“ 13 = „Ich beschäftige mich damit, was noch passieren könnte.“ 17 = „Ich mache mir Sorgen, dass andere schlecht über mich denken.“ 21 = „Ich bin besorgt, dass andere besser sind als ich.“ 25 = „Wenn mir etwas Unangenehmes passiert ist, gehen mir die Gedanken daran nicht aus dem Kopf.“

TR: 4 = „Ich habe das Gefühl, dass ich weinen möchte.“ 8 = „Ich bin traurig.“ 14 = „Ich fühle mich allein, auch wenn ich mit anderen zusammen bin.“ 18 = „Ich bin unglücklich.“ 26 = „Ich bin einsam.“

FR: 1 = „Ich bin fröhlich.“ 6 = „Mir geht es gut.“ 11 = „Ich bin glücklich.“ 15 = „Ich finde das Leben schön.“ 19 = „Ich bin gut gelaunt.“ 23 = „Ich bin gut drauf.“ 27 = „Ich freue mich.“

### Skalengkorrlationen, Reliabilitäten und Stabilitäten

Die Korrelationen der Skalen variierten betragsmäßig zwischen  $|r| = .29$  und  $.57$ , ihr Muster entspricht den Faktorkorrelationen (Tabelle 1). Die Reliabilitäten streuten zwischen  $\alpha = .70$  und  $.88$  und fielen für die Aufgeregtheit etwas niedriger aus als für die anderen Skalen (Aufgeregtheit:  $.70$ , Besorgnis:  $.78$ , Traurigkeit:  $.78$ , Freude:  $.88$ ). Für die beiden zusammengefassten Skalen (Angst, Depression), deren Korrelation  $.50$  betrug, ergaben sich Werte von  $\alpha = .83$  (Angst) und  $.88$  (Depression).

Die Stabilitäten wurden anhand einer Teilstichprobe von 497 Kindern der Klassenstufen 3 bis 10 (Grundschule: 34 %, Gymnasium: 39 %, Realschule: 27 %) geprüft. Ein Teil der Testungen fand im Abstand von zwei Wochen, ein anderer Teil im Abstand von vier Wochen statt. Die Stabilitäten fielen für beide Zeitabstände nahezu identisch aus. Die Stabilitäten waren durchweg geringer als die Reliabilitäten (Angst, 2 Wochen / 4 Wochen: .72 / .74, Aufgeregtheit: .61 / .61, Besorgnis: .71 / .73, Depression: .76 / .76, Traurigkeit: .68 / .69, Freude: .74 / .72), was impliziert, dass die Testergebnisse zu einem gewissen Maße von variablen situativen Einflüssen abhängen.

### **Messinvarianz**

Fitindices für die zur Prüfung der Messinvarianz berechneten Mehrgruppenmodelle sind in Tabelle 4 zu entnehmen. Wie erwartet, sind die  $\chi^2$ -Werte durchweg hochsignifikant und nehmen bei restriktiveren Modellen jeweils deutlich zu. Eine Ausnahme betrifft hier den Vergleich konfiguraler und schwacher Invarianz – hier nehmen die  $\chi^2$ -Werte ab. Die Abnahme ist darin begründet, dass die  $\chi^2$ -Werte bei der WLSMV-Schätzung modellspezifisch skaliert werden, weshalb sie auch bei verschachtelten Modellen (wie im Fall konfiguraler vs. schwacher Invarianz) bei restriktiveren gegenüber wenig restriktiveren Modellen nicht unbedingt zunehmen müssen (und daher auch nicht gegeneinander getestet werden sollten; siehe Satorra & Bentler, 2010). Gemessen an den approximativen Fitindices weisen jedoch alle Modelle akzeptable Passungen auf (strikte Invarianz für Geschlechtsgruppen: CFI = .961, RMSEA = .044; für Altersgruppen: CFI = .954, RMSEA = .047). Die Veränderungen im CFI und RMSEA sind durchweg relativ gering (kleiner als .01), bei der schwachen gegenüber der konfiguralen Invarianz resultieren sogar leichte Verbesserungen der Passung. Struktur und Messeigenschaften des MAI-KJ scheinen bei Jungen und Mädchen sowie bei den drei Altersgruppen also insgesamt ähnlich zu sein. Die strikteren Kriterien von Meade et al. (2008) wurden allerdings hinsichtlich der starken Invarianz nicht vollständig erfüllt, da sich hier beim CFI Veränderungen zeigten, die größer als .002 sind.

### **Zusammenhänge mit anderen Verfahren**

Die Korrelationen der MAI-KJ-Skalen mit Skalen aus anderen Verfahren sind in Tabelle 5 tabellarisch zusammengestellt. Hier finden sich auch *t*-Tests für abhängige Korrelationen

(Steiger, 1980), mit denen geprüft wurde, ob sich die Korrelationen der jeweiligen Skala mit den Angst- und Depressionsgesamtscores des MAI-KJ bedeutsam unterscheiden. Diese Prüfung ist im Hinblick auf die diskriminante Validität der MAI-KJ-Skalen von Interesse: Wie in ESM 1 im Einzelnen dargestellt, erwarten wir, dass die Angstskalen anderer Verfahren höher mit der Angstskala als mit der Depressionsskala des MAI-KJ korrelieren; entsprechend sollen die Depressionsskalen anderer Verfahren höher mit der Depressionsskala als mit der Angstskala des MAI-KJ korrelieren.

Wir konzentrieren uns hier auf die Darstellung der Korrelationen mit konstruktnahen Verfahren zur Messung von Angst und Depression (*Angstfragebogen für Schüler*, AFS; Wieczerkowski, Nickel, Janowski, Fittkau & Rauer, 1981; *Kinder-Angst-Test-II*, KAT-II; Thurner & Tewes, 2000; *Social Anxiety Scale for Children – revised*, deutsche Version, SASC-R-D; Melfsen, 1998; *Depressionstest für Kinder*, DTK-II; Rossmann, 2014; *Depressionstest für Kinder im Grundschulalter*, DTGA; Esser, Laucht, Drews & Ihle, 2013).

Für die Angstskalen des AFS und des KAT II zeigten sich erwartungskonform hohe positive Zusammenhänge mit den Angstskalen des MAI-KJ ( $r_s$  zwischen .52 und .71, Aufgeregtheit:  $r_s$  zwischen .42 und .61, Besorgnis:  $r_s$  zwischen .49 und .66), die stärker ausgeprägt waren als die Korrelationen mit den MAI-KJ-Depressionsskalen (Depression:  $r_s$  zwischen .31 und .59,  $t \geq 3.32$ ,  $p < .05$ , Traurigkeit:  $r_s$  zwischen .34 und .62, Freude:  $r_s$  zwischen  $-.23$  und  $-.46$ ). Ebenfalls den Erwartungen entsprechend war die Skala Schulunlust des AFS stärker mit Depression ( $r = .40$ ) als mit Angst ( $r = .24$ ,  $t = 3.84$ ,  $p < .05$ ) verbunden. Die im AFS zusätzlich erfasste Tendenz, sozial erwünscht zu antworten, korrelierte schwach negativ mit der MAI-KJ-Depression ( $r = -.24$ ). Für die Angstskalen des MAI-KJ und die Subskala Traurigkeit zeigten sich Korrelationen nahe Null mit sozialer Erwünschtheit.

Die mittels des SASC-R-D erfassten sozialen Ängste waren ebenfalls positiv mit den Angstskalen des MAI-KJ assoziiert ( $r_s$  zwischen .35 und .60). Die höchste Korrelation zeigte sich hier zwischen der MAI-KJ-Besorgnis und der Subskala Furcht vor negativer Bewertung ( $r = .60$ ). Ferner waren soziale Ängste des SASC-R-D deutlich mit den MAI-KJ-Depressionsskalen (insbesondere Traurigkeit) korreliert, allerdings weniger stark als mit der MAI-KJ-Angst,  $|r|$  zwischen .25 und .51,  $t \geq 2.91$ ,  $p < .05$ .

Es zeigten sich substantielle Zusammenhänge der Subskala Dysphorie und Selbstwertprobleme des DTK-II sowie des Gesamtwerts Depression des DTGA mit den Depressionsskalen des MAI-KJ,  $|r|$  zwischen .41 und .63. Etwas niedriger fielen die Korrelationen der Subskalen Agitiertes Verhalten und Müdigkeit/autonome Reaktionen des DTK-II mit den Depressionsskalen des MAI-KJ aus,  $|r|$  zwischen .25 und .38. Korrelationen in jeweils ähnlicher Höhe zeigten sich mit den Angstskalen des MAI-KJ (Dysphorie und Selbstwertprobleme:  $r_s$  zwischen .37 und .59, Agitiertes Verhalten sowie Müdigkeit / autonome Reaktionen:  $r_s$  zwischen .21 und .41). Signifikante Korrelationsdifferenzen zwischen der MAI-KJ-Angst- und Depressionsskala ließen sich hier nicht aufweisen.

### **Diskussion**

Vorgestellt wurde das allgemeine Modul des MAI-KJ (Eschenbeck et al., in Vorb.). Dieses Modul ist als Screening-Instrument konzipiert und soll situationsübergreifende Dispositionen zur Manifestation von Angst (Aufgeregtheit, Besorgnis) und Depression (Traurigkeit, Anhedonie / Mangel an Freude) erfassen. Die erwartete Struktur des Verfahrens konnte mit Hilfe explorativer und darauf aufbauender konfirmatorischer Faktorenanalysen an zwei großen, voneinander unabhängigen Stichproben bestätigt werden. Dabei wurden die Items eindeutig einem Besorgnis-, Aufgeregtheits-, Traurigkeits- bzw. Freundefaktor zugeordnet. Die Korrelationen zwischen den Faktoren bzw. Skalen waren substantiell, jedoch nicht so hoch, dass deren Eigenständigkeit in Frage gestellt würde. Das Korrelationsmuster der Variablen und die gute Passung eines Modells mit übergeordneten Angst- bzw. Depressionsfaktoren deutet darauf hin, dass auch eine Betrachtung stärker aggregierter Angst- bzw. Depressionsscores möglich ist, was für bestimmte Forschungs- wie auch praktische Fragen zweckmäßig sein kann.

Die Resultate untermauern die Separierbarkeit der mit dem MAI-KJ erfassten Angst- und Depressionsfacetten bei Kindern und Jugendlichen. Traurigkeit ist allerdings sehr hoch mit den beiden Angstfaktoren (Besorgnis und Aufgeregtheit) assoziiert (latente Korrelation jeweils .72). Dies findet sich in ähnlicher Weise beim STADI (siehe Renner et al., 2018, Tabelle 1). Die hohen Korrelationen gehen also nicht etwa auf eine geringere Differenziertheit des Emotionsberichts bei Kindern und Jugendlichen im Vergleich zu Erwachsenen zurück. Sie sind

überdies im Einklang mit anderen Befunden zu Assoziationen zwischen positiven und negativen Affekten (siehe z. B. Watson et al., 1988). In der Anwendung muss also mit deutlichen Niveauunterschieden zwischen Personen gerechnet werden, die mehrere Skalen betreffen (z. B. wird erhöhte Traurigkeit häufig mit erhöhter Besorgnis und/oder Aufgeregtheit einhergehen). Die Korrelationen zwischen den Faktoren bzw. Skalen lassen dennoch viel Raum für genuine Profildifferenzen (z. B. erhöhte Besorgnis bei durchschnittlich ausgeprägter Traurigkeit), die für die Planung problemspezifischer Interventionen besonders informativ sind.

Die Prüfung der Messinvarianz bezüglich Geschlecht und Alter indizierte, dass zumindest schwache Invarianz, welche die Faktorenstruktur und die Faktorladungen betrifft, gegeben ist. Im Hinblick auf stärkere Formen der Invarianz ist die Evidenz mehrdeutig, da die Bewertung der Ergebnisse von den jeweils verwendeten Kriterien abhängt, die – wie angedeutet wurde – bei Faktorenmodellen für kategoriale Variablen zurzeit noch auf einer unsicheren Basis stehen (Sass et al., 2014). Die Tatsache, dass auch Modelle mit strikten Invarianzanforderungen eine (gemessen an approximativen Fitindices) gute Passung aufwiesen, spricht allerdings dafür, dass bei Jungen und Mädchen sowie in diversen Altersgruppen im Wesentlichen die gleichen Merkmale mit der gleichen Genauigkeit abgebildet werden. Dies ist für alle Forschungsfragen wichtig, in denen Gruppenvergleiche eine Rolle spielen. Bei mangelnder Messinvarianz können Ähnlichkeiten und Unterschiede zwischen Gruppen nicht sinnvoll interpretiert werden, da sich die Testscores dann in jeder Gruppe auf andere Eigenschaften beziehen. Zur Invarianz sind weitere Untersuchungen sinnvoll, in denen neben Alter und Geschlecht auch andere potenziell relevante Gruppierungsmerkmale einbezogen werden könnten. Dabei wäre es interessant, den praktischen Implikationen mangelnder oder nur partiell vorliegender Invarianz nachzugehen (z. B. für die Berechnung von Itemsammen- oder Itemmittelwerten als Indikatoren für die zu messenden Eigenschaften). Hierfür bietet es sich an, auf neuere Entwicklungen im Rahmen der Item-Response-Theorie zurückzugreifen, in denen Messinvarianz (auch) auf der in Anwendungen primär interessierenden Testebene (*differential test functioning*) betrachtet werden kann (Chalmers, Counsell & Flora, 2016).

Die Reliabilitäten (geschätzt über Cronbachs  $\alpha$ ) lagen für Angst, Depression und Freude über der anvisierten Marke von .80. Für die Subskalen Besorgnis und Traurigkeit fielen sie mit  $\alpha = .78$  knapp geringer aus. Damit sind diese Skalen auch für die Einzelfalldiagnostik gut verwendbar. Lediglich für die Subskala Aufgeregtheit war die Reliabilität mit  $\alpha = .70$  deutlich niedriger als erwünscht. Die Reliabilität der Aufgeregtheitsskala ist für Forschungszwecke zwar ausreichend, allerdings muss bei der Einzelfalldiagnostik die im Vergleich zu den anderen Skalen größere Ungenauigkeit der Messwerte in der Interpretation berücksichtigt werden.

Im Vergleich hierzu fielen die für einen Zeitabstand von zwei bis vier Wochen bestimmten Stabilitäten etwas niedriger aus (zwischen .61 und .76), was einen gewissen situativen Einfluss auf die Messung belegt. Die Werte sind aber hoch genug, um von der Messung relativ stabiler Dispositionen ausgehen zu können. In künftigen Untersuchungen wäre es interessant, den relativen Anteil stabiler und situativer Einflüsse auf die Messwerte genauer, z. B. mit Latent-State-Trait-Analysen (Steyer, Mayer, Geiser & Cole, 2015), zu prüfen.

Analysen der Zusammenhänge der MAI-KJ-Skalen mit korrespondierenden Skalen aus anderen Verfahren zur Diagnostik von Angst und Depression (soziale Ängstlichkeit, allgemeine Ängstlichkeit, diverse depressionsbezogene Merkmale) erbrachten für die Angstskalen des MAI-KJ deutliche Konvergenzen: Die Korrelationen mit Angstskalen aus anderen Verfahren waren substanziell (mittleres  $r = .58$ ) und durchweg höher als die entsprechenden Korrelationen der MAI-KJ-Depressionsskalen (mittleres  $r = .44$ ).

Auch für die Depressionsskalen des MAI-KJ zeigten sich deutliche konvergente Beziehungen zu den entsprechenden Skalen aus anderen Verfahren (mittleres  $r = .47$ ). Allerdings waren die Zusammenhänge der MAI-KJ-Angstskalen mit den Depressionsskalen aus anderen Verfahren nahezu ebenso hoch ausgeprägt (mittleres  $r = .44$ ). Dies entspricht bisherigen Befunden: So berichtet z. B. Rossmann (2014) für die Skala Dysphorie / Selbstwertprobleme des DTK von Korrelationen mit Angstmaßen zwischen  $r = .40$  und  $.49$ . Die hohen Korrelationen zwischen Depressions- und Angstmaßen dürften zum Teil (wenn auch sicherlich nicht vollständig; siehe Renner et al., 2018) darauf zurückgehen, dass in Depressionstests im Allgemeinen auch Items verwendet werden, die inhaltlich Angst, und hier

insbesondere Besorgnis, ansprechen (im DTK z. B. „Hast du oft Angst, etwas falsch zu machen?“ oder „Hast du mehr Sorgen als andere Kinder?“).

Insgesamt stützen die Resultate Erwartungen an die Konstruktvalidität des MAI-KJ; die wenigen nicht antizipierten Befunde sind durchweg psychologisch sinnvoll interpretierbar. Einige Limitationen, welche die Messinvarianz sowie den relativen Einfluss stabiler und variabler Faktoren auf die Antworten betreffen, wurden benannt. Zudem sind insbesondere zur Validierung noch umfassendere Studien notwendig, in denen z. B. auch objektive Indikatoren der Angst und Depression berücksichtigt oder Selbst- und Fremdeinschätzungen (etwa die Sicht der Eltern, Lehrkräfte oder Peers auf emotionale Probleme des Kindes) miteinander kontrastiert werden. Auch eine Einbeziehung von Kindern und Jugendlichen mit klinisch relevanten Problemen wäre sicherlich vielversprechend. Die dargestellten Ergebnisse weisen den allgemeinen Teil des MAI-KJ als ein valides und ökonomisches Instrument zur simultanen Erfassung von Angst und Depression bei Kindern und Jugendlichen im Altersbereich von 8 bis 16 Jahren aus. Vom weiten Altersbereich des MAI-KJ können insbesondere Studien zur Veränderung von Angst und Depression im Kindes- und Jugendalter profitieren.

Das MAI-KJ kann immer dann eingesetzt werden, wenn eine differenzierte Erfassung angst- bzw. depressionsbezogener Befindlichkeiten für Forschungs- oder praktische Fragen erwünscht ist. Es ist das einzige deutschsprachige Instrument, das eine diskriminant valide Diagnose von Angst *und* Depression bei Kindern und Jugendlichen ermöglicht. Im Unterschied zu anderen Verfahren in diesem Bereich ist das MAI-KJ mit einem adäquaten Messmodell abgesichert und basiert auf aktuellen Standards entsprechenden Methoden. Das MAI-KJ ist zudem auch das einzige Verfahren, für das Messinvarianz in Bezug auf Alter und Geschlecht geprüft und weitgehend nachgewiesen wurde. Dies ist vor allem für Forschungszwecke essenziell, um Ergebnisse aus verschiedenen Gruppen sinnvoll miteinander vergleichen zu können. Besonders Forschung zu Gemeinsamkeiten und Unterschieden von Angst und Depression kann vom Einsatz des MAI-KJ profitieren. Bisher konzentriert sich diese Forschung vornehmlich auf Erwachsene. Das MAI-KJ erleichtert es nun, auch Kinder und Jugendliche verstärkt in die Forschung einzubeziehen.

### Literatur

- Anderson E. R. & Hope D. A. (2008). A review of the tripartite model for understanding the link between anxiety and depression in youth. *Clinical Psychology Review, 28*, 275–287. doi:10.1016/j.cpr.2007.05.004
- American Psychiatric Association (APA; 2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed., DSM-5). Washington, DC: Author.
- Brodbeck, J., Abbott, R. A., Goodyer, I. M. & Croudace, T. J. (2011). General and specific components of depression and anxiety in an adolescent population. *BMC Psychiatry, 11*:191. doi:10.1186/1471-244X-11-191
- Beauducel, A. & Wittmann, W. W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling, 12*, 41–75. doi:10.1207/s15328007sem1201\_3
- Chalmers, R. P., Counsell, A. & Flora, D. B. (2016). It might not make a big DIF: Improved differential test functioning statistics that account for sampling variability. *Educational and Psychological Measurement, 76*, 114–140. doi:10.1177/0013164415584576
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. doi:10.1080/10705510701301834
- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. doi:10.1207/S15328007SEM0902\_5
- Clark, L. A. & Watson, D. (1991). Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*, 316–336. doi:10.1037/0021-843X.100.3.316
- Costa, P., Jr., Terracciano, A. & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*, 322–331. doi:10.1037//0022-3514.81.2.322
- Cummings, C. M., Caporino, N. E. & Kendall, P. C. (2014). Comorbidity of anxiety and depression in children and adolescents: 20 years after. *Psychological Bulletin, 140*, 816–845. doi:10.1037/a0034733

- Eschenbeck, H., Heim-Dreger, U., Hock, M. & Kohlmann, C.-W. (in Vorb.). *Mehrdimensionales Angstinventar für Kinder und Jugendliche (MAI-KJ)*. Göttingen: Hogrefe.
- Esser, G., Laucht, M., Drews, S. & Ihle, W. (2013). *Depressionstest für Kinder im Grundschulalter (DTGA)*. Göttingen: Hogrefe.
- Fehm, L. & Fydrich, T. (2011). *Prüfungsangst*. Göttingen: Hogrefe.
- Hofmann, R. J. (1978). Complexity and simplicity as objective indices descriptive of factor solutions. *Multivariate Behavioral Research*, 13, 247–250. doi:10.1207/s15327906mbr1302\_9
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Krohne, H. W. (2010). *Psychologie der Angst*. Stuttgart: Kohlhammer.
- Laux, L., Hock, M., Bergner-Köther, R., Hodapp, V. & Renner, K.-H. (2013). *Das State-Trait-Angst-Depressions-Inventar (STADI)*. Göttingen: Hogrefe.
- Liebert, R. M. & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 20, 975–978. doi:10.2466/pr0.1967.20.3.975
- Lonigan, C. J. & Phillips, B. M. & Hooe, E. S. (2003). Relations of positive and negative affectivity to anxiety and depression in children: Evidence from a latent variable longitudinal study. *Journal of Consulting and Clinical Psychology*, 71, 465–481. doi:10.1037/0022-006X.71.3.465
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- 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, 568–592. doi:10.1037/0021-9010.93.3.568
- Melfsen, S. (1998). Die deutsche Fassung der Social Anxiety Scale for Children Revised (SASC-R-D): Psychometrische Eigenschaften und Normierung. *Diagnostica*, 44, 1–11.

- Millsap, R. E. & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, 39, 479–515. doi:10.1207/S15327906MBR3903\_4
- Mineka, S., Watson, D. & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology*, 49, 377–412. doi:10.1146/annurev.psych.49.1.377
- Muthén, B. O. & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus*. Zugriff am 2.4.2017 unter <https://www.statmodel.com/download/webnotes/CatMGLong.pdf>
- Nolen-Hoeksema, S., Wisco, B. E. & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, 3(5), 400–424. doi:10.1111/j.1745-6924.2008.00088.x
- Plück, J., Döpfner, M. & Lehmkuhl, G. (2000). Internalisierende Auffälligkeiten bei Kindern und Jugendlichen in Deutschland – Ergebnisse der PAK-KID-Studie. *Kindheit und Entwicklung*, 9, 133–142. doi:10.1026//0942-5403.9.3.133
- Ravens-Sieberer, U., Wille, N., Bettge, S. & Erhart, M. (2007). Psychische Gesundheit von Kindern und Jugendlichen in Deutschland: Ergebnisse aus der BELLA-Studie im Kinder- und Jugendgesundheitsurvey (KiGGS). *Bundesgesundheitsblatt*, 50, 871–878. doi:10.1007/s00103-007-0250-6
- Renner, K.-H., Hock, M., Bergner-Köther, R. & Laux, L. (2018). Differentiating anxiety and depression: The State-Trait Anxiety-Depression Inventory. *Cognition and Emotion*, 32(7), 1409–1423. doi:10.1080/02699931.2016.1266306
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. URL: <http://www.jstatsoft.org/v48/i02/>
- Rossmann, P. (2014). *Depressionstest für Kinder (DTK-II)*. Göttingen: Hogrefe.
- Rost, D. H. & Haferkamp, W. (1979). Zur Brauchbarkeit des AFS (Angstfragebogen für Schüler). *Zeitschrift für Empirische Pädagogik*, 3, 183–210.
- Sass, D. A., Schmitt, T. A. & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators.

- Structural Equation Modeling: A Multidisciplinary Journal*, 21, 167–180.  
doi:10.1080/10705511.2014.882658
- Satorra, A. & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75, 243–248. doi:10.1007/s11336-009-9135-y
- Spence, S. H. (2017). Review – Measurement Issues: Assessing anxiety disorders in children and adolescents. *Child and Adolescent Mental Health*, 1–17. doi:10.1111/camh.12251
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19, 85–95.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251. doi:10.1037/0033-2909.87.2.245
- Steyer, R., Mayer, A., Geiser, C. & Cole, D. A. (2015). A theory of states and traits - revised. *Annual Review of Clinical Psychology*, 11, 71–98. doi:10.1146/annurev-clinpsy-032813-153719
- Turner, F. & Tewes, U. (2000). *Kinder-Angst-Test-II (KAT II)*. Göttingen: Hogrefe.
- Watson, D. (2009). Differentiating the mood and anxiety disorders: A quadripartite model. *Annual Review of Clinical Psychology*, 5, 221–247. doi:10.1146/annurev.clinpsy.032408.153510
- Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. doi:10.1037/0022-3514.54.6.1063
- World Health Organization (WHO; 1993). *The ICD-10 classification of mental and behavioural disorders. Diagnostic criteria for research* (10th rev., ICD-10). Genf, Schweiz: Author.
- Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B. & Rauer, W. (1974). *Angstfragebogen für Schüler (AFS)* (1. Aufl.). Göttingen: Hogrefe.
- Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B. & Rauer, W. (1981). *Angstfragebogen für Schüler (AFS)* (2. Aufl.). Göttingen: Hogrefe.

- Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B., Rauer, W. & Petermann, F. (2016). *Angstfragebogen für Schüler (AFS)* (7. Aufl.). Göttingen: Hogrefe.
- Zavaglia, E. & Bergeron, L. (2017). Systematic review of comorbidity between DSM disorders and depression according to age and sex in youth. *Canadian Psychology / Psychologie canadienne*, 58, 124–139. doi:10.1037/cap0000085.
- Zeidner, M. (1998). *Test anxiety: The state of the art*. New York, NY: Plenum Press.

## **Elektronische Supplemente**

### ESM 1. Verfahren zur Validierung des MAI-KJ

Hier werden die zur Validierung des MAI-KJ eingesetzten Verfahren und Stichproben sowie die Erwartungen beschrieben.

### ESM 2. Stichproben

Hier werden die Stichproben, die Itemgewinnung und die Behandlung fehlender Werte beschrieben.

### ESM 3. Explorative Faktorenanalysen

Hier werden die explorativen Faktorenanalysen des MAI-KJ beschrieben.

### ESM 4. Messinvarianz

Hier wird die Prüfung der Messinvarianz des MAI-KJ beschrieben.

### ESM 5. Validierungskorrelationen

Hier werden die Korrelationen des MAI-KJ mit anderen Verfahren beschrieben und diskutiert.

*Elektronisches Supplement 1. Verfahren zur Validierung des MAI-KJ*

Eine Übersicht der zur Validierung des MAI-KJ eingesetzten Verfahren, der Stichproben und der Erwartungen findet sich in Tabelle 1.

*Angstfragebogen für Schüler (AFS; Wieczerkowski, Nickel, Janowski, Fittkau & Rauer, 1981; Wieczerkowski et al, 2016).* Der AFS enthält vier Subskalen (zweistufiges Antwortformat: *stimmt* = 1 oder *stimmt nicht* = 0): Prüfungsangst (15 Items, z. B. „Vor Klassenarbeiten bin ich immer aufgeregt“, „Wenn eine Klassenarbeit geschrieben wird, mache ich oft Fehler, weil ich zu viel Angst habe“), manifeste Angst (15 Items, z. B. „Oft muss ich daran denken, dass mir etwas zustoßen könnte“, „Ich werde oft ganz nervös“), Schulunlust (10 Items, z. B. „Es wäre schön, wenn ich nicht mehr zur Schule zu gehen brauchte“), soziale Erwünschtheit (10 Items, z. B. „Ich sage immer die Wahrheit“). Wir erwarteten positive Korrelationen zwischen den Angstskalen des MAI-KJ und des AFS, die stärker ausgeprägt sein sollten als die Korrelationen der Depressionsskalen des MAI-KJ mit den Angstskalen des AFS. Dagegen erwarteten wir für die Skala Schulunlust des AFS stärkere Zusammenhänge mit den Depressionsskalen des MAI-KJ als mit den Angstskalen. Die AFS-Subskalen erwiesen sich in der Analysestichprobe als reliabel (PA: Cronbachs  $\alpha = .86$ , MA:  $\alpha = .84$ , SU:  $\alpha = .73$ , SE:  $\alpha = .71$ ).

*Kinder-Angst-Test-II (KAT-II; Thurner & Tewes, 2000).* Form A des KAT-II erfasst mit 18 Items (zweistufiges Antwortformat: *ja* = 1 oder *nein* = 0, z. B. „Ich leide oft unter Kopfschmerzen“, „Ich mache mir manchmal Sorgen um die Zukunft“) die Ängstlichkeit des Kindes. Wiederum wurden positive Zusammenhänge für die MAI-KJ-Angstskalen erwartet, die stärker ausgeprägt sein sollten als für die Depressionsskalen. In der Analysestichprobe betrug die interne Konsistenz  $\alpha = .86$ .

*Social Anxiety Scale for Children – revised (deutsche Version, SASC-R-D; Melfsen, 1998).* Die SASC-R-D zur Erfassung sozialer Ängste von Kindern und Jugendlichen besteht aus zwei Subskalen mit je neun Items (fünfstufiges Antwortformat: *nie* = 1 bis *immer* = 5). Die Skala *Fear of Negative Evaluation* (FNE) beinhaltet Gedanken zur Anerkennung bzw. Ablehnung durch andere Personen (z. B. „Ich glaube, dass andere Jungen und Mädchen hinter meinem Rücken über mich reden“), die Skala *Social Avoidance and Distress* (SAD) benennt Situationen, die Angst auslösen bzw. die vermieden werden (z. B. „Ich fühle mich unsicher, selbst bei Jungen und Mädchen, die ich sehr gut kenne“). Erwartet wurden positive Zusammenhänge für die MAI-KJ-Angstskalen, die stärker ausgeprägt sein sollten als die Korrelationen mit den MAI-KJ-Depressionsskalen. Insbesondere die Korrelation zwischen Besorgnis und FNE sollte hoch ausfallen, da beide Skalen kognitive Angstreaktionen thematisieren. In der Analysestichprobe betrug die internen Konsistenzen  $\alpha = .90$  für FNE und  $\alpha = .80$  für SAD.

*Depressionstest für Kinder (DTK-II; Rossmann, 2014).* Die Langform des DTK-II erfasst mit 55 Items (zweistufiges Antwortformat: *nein* = 0, *ja* = 1) die depressive Befindlichkeit von Kindern mit Hilfe der drei Subskalen *Dysphorie / Selbstwertprobleme* (25 Items, z. B. „Fühlst du dich oft wertlos?“), *agitiertes Verhalten* (16 Items, z. B. „Gibst du manchmal in der Schule freche Antworten?“) und *Müdigkeit / autonome Reaktionen* (14 Items, z. B. „Bist du oft den ganzen Tag müde?“). In der vorliegenden Studie wurde die Subskala *Dysphorie / Selbstwertprobleme* um ein Item („Möchtest du manchmal tot sein?“) gekürzt, da dieses Item für die Befragung in den Schulklassen als ungeeignet betrachtet wurde. Hohe Korrelationen wurden besonders für die depressionsbezogenen MAI-KJ-Skalen (*Traurigkeit*, *Freunde*) mit der DTK-II-Subskala *Dysphorie / Selbstwertprobleme* erwartet, da diese Skala affektive und kognitive Symptomausprägungen fokussiert. Darüber hinaus sollten die Zusammenhänge mit den MAI-KJ-Depressionsskalen stärker ausgeprägt sein als mit den Angstskalen. In der Analysestichprobe variierten die internen Konsistenzen zwischen  $\alpha = .90$  für *Dysphorie / Selbstwertprobleme* und  $\alpha = .78$  für *Agitiertes Verhalten*.

*Depressionstest für Kinder im Grundschulalter (DTGA; Esser, Laucht, Drews & Ihle, 2013).* Der DTGA erfasst mit 12 Items (dreistufiges Antwortformat: *stimmt gar nicht* = 0 bis *stimmt ganz genau* = 2, z. B. „Ich bin oft traurig“) das Ausmaß depressiver Symptomatik von Kindern. Auch hier sollten die Zusammenhänge mit den depressionsbezogenen MAI-KJ-Skalen stärker ausgeprägt sein als die Zusammenhänge mit den Angstskalen. In der Analysestichprobe betrug die interne Konsistenz  $\alpha = .68$ .

Tabelle 1. Zur Validierung eingesetzte Verfahren

Verfahren	Subskalen und Reliabilitäten	Stichprobe	Erwartungen
Angstfragebogen für Schüler (AFS; Wiczerkowski, Nickel, Janowski, Fittkau & Rauer, 1981)	Prüfungsangst (15 Items, $\alpha = .86$ ) Manifeste Angst (15 Items, $\alpha = .84$ ) Schulunlust (10 Items, $\alpha = .73$ ) Soziale Erwünschtheit (10 Items, $\alpha = .71$ )	$N = 433$ , 43 % Mädchen, Klassenstufen 3–10, Alter: $M = 12.09$ , $SD = 2.29$	Angstskalen: $A > D$ Schulunlust: $D > A$
Kinder-Angst-Test-II (KAT-II, Form A; Thurner & Tewes, 2000)	Dispositionelle Ängstlichkeit (18 Items, $\alpha = .86$ )	$N = 325$ , 44 % Mädchen, Klassenstufen 5–10, Alter: $M = 12.89$ , $SD = 1.65$	$A > D$
Social Anxiety Scale for Children – revised (deutsche Version, SASC-R-D; Melfsen, 1998)	Fear of Negative Evaluation (9 Items, $\alpha = .90$ ) Social Avoidance and Distress (9 Items, $\alpha = .80$ )	$N = 530$ , 55 % Mädchen, Klassenstufen 5–10, Alter: $M = 13.07$ , $SD = 1.84$	$A > D$
Depressionstest für Kinder (DTK-II; Rossmann, 2014)	Dysphorie / Selbstwertprobleme (25 Items, $\alpha = .90$ ) Agitiertes Verhalten (16 Items, $\alpha = .78$ ) Müdigkeit / autonome Reaktionen (14 Items, $\alpha = .81$ )	$N = 486$ , 48 % Mädchen, Klassenstufen 3–7, Alter: $M = 10.34$ , $SD = 1.44$	Besonders für Dysphorie: $D > A$
Depressionstest für Kinder im Grundschulalter (DTGA; Esser, Laucht, Drews & Ihle, 2013)	Depressive Symptomatik (12 Items, $\alpha = .68$ )	$N = 109$ , 45 % Mädchen, Klassenstufen 3–4, Alter: $M = 9.55$ , $SD = 0.73$	$D > A$

*Anmerkungen:* Erwartungen bezüglich der Korrelationen der aufgeführten Skalen mit den MAI-KJ-Skalen: + / -A, + / -D = positive / negative Korrelationen mit den Angst- (A) bzw. Depressionsskalen (D) des MAI-KJ.  $A > D$  = Die Korrelationen mit den Angstskalen des MAI-KJ sollte höher ausfallen als mit den depressionsbezogenen Skalen des MAI-KJ.  $D > A$  = umgekehrtes Muster.

### Literaturverzeichnis

- Esser, G., Laucht, M., S. Drews & W. Ihle (2013). *Depressionstest für Kinder im Grundschulalter (DTGA)*. Göttingen: Hogrefe.
- Melfsen, S. (1998). Die deutsche Fassung der Social Anxiety Scale for Children Revised (SASC-R-D): Psychometrische Eigenschaften und Normierung. *Diagnostica*, 44, 1-11. doi:10.1026//0012-1924.45.2.95
- Rossmann, P. (2014). *Depressionstest für Kinder (DTK-II)*. Göttingen: Hogrefe.
- Thurner, F. & Tewes, U. (2000). *Kinder-Angst-Test-II (KAT II)*. Göttingen: Hogrefe.

- Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B. & Rauer, W. (1981). *Angstfragebogen für Schüler (AFS)* (2. Aufl.). Göttingen: Hogrefe.
- Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B., Rauer, W. & Petermann, F. (2016). *Angstfragebogen für Schüler (AFS)* (7. Aufl.). Göttingen: Hogrefe.

*Elektronisches Supplement 2. Stichproben, Itemgewinnung und Behandlung fehlender Werte*

*Stichproben.* Stichprobe 1, anhand derer der explorativen Faktorenanalysen durchgeführt wurden, umfasste 135 Klassen aus 37 Schulen (Grundschule: 30 %, Hauptschule: 11 %, Realschule: 30 %, Gymnasium: 26 %, andere Schularten: 3 %). Stichprobe 2, die für die konfirmatorischen Faktorenanalysen verwendet wurde, umfasste 412 Klassen aus 97 Schulen (Grundschule: 35 %, Hauptschule: 17 %, Realschule: 30 %, Gymnasium: 18 %). Die Verteilung auf die Klassenstufen 3 bis 10 betrug für die erste Stichprobe 17 % (Stufe 3), 14 % (4), 15 % (5), 11 % (6), 20 % (7), 13 % (8), 5 % (9) und 4 % (10), für die zweite Stichprobe 15 % (3), 20 % (4), 13 % (5), 16 % (6), 10 % (7), 12 % (8), 9 % (9) und 6 % (10). Die beiden höchsten Klassenstufen waren in den Stichproben also leicht unterrepräsentiert. 36 % (Stichprobe 1) bzw. 33 % (Stichprobe 2) der Kinder gaben an, ihre Mutter oder ihr Vater stamme nicht aus Deutschland (keiner der beiden Elternteile aus Deutschland: 22 % bzw. 18 %).

Mittelwerts- und Streuungsunterschiede der MAI-KJ-Skalen zwischen den Bundesländern (Baden-Württemberg, Bayern, Nordrhein-Westfalen), aus denen die Stichproben rekrutiert worden waren, fielen, wie Tabelle 1 für die Endversion des Verfahrens (Stichprobe 2) zu entnehmen ist, nur sehr gering aus. Für die Analyse der Unterschiede wurde für jede MAI-KJ-Skala ein Mehrebenenmodell mit Bundesland (dreistufig) als fixem Faktor und der Klassenzugehörigkeit als Zufallsfaktor berechnet. Als Kovariaten gingen Geschlecht, Schultyp (Grundschule, Hauptschule, Realschule, Gymnasium) sowie die Klassenstufe ein. Die Analysen wurden mit dem R-Paket *lme4* durchgeführt (Bates, Maechler, Bolker & Walker, 2015). In keiner der vier Analysen erwies sich das Bundesland als statistisch bedeutsamer Prädiktor.

*Tabelle 1.* Item-Mittelwerte und Standardabweichungen (in Klammern) der MAI-KJ-Skalen in Baden-Württemberg, Bayern und Nordrhein-Westfalen für die Endversion des Verfahrens

<b>Skala</b>	<b>Baden-Württemberg</b>		<b>Bayern</b>		<b>Nordrhein-Westfalen</b>	
Aufgeregtheit	1.75	(0.48)	1.76	(0.48)	1.68	(0.50)
Besorgnis	2.02	(0.58)	1.96	(0.59)	1.93	(0.58)
Traurigkeit	1.54	(0.53)	1.55	(0.53)	1.49	(0.50)
Freude	3.23	(0.58)	3.27	(0.56)	3.33	(0.59)

*Anmerkungen:* Die Werte basieren auf Stichprobe 2. Die Skalenwerte reichen von 1 bis 4. Die Stichprobenumfänge betragen  $n = 5\,990$  (Baden-Württemberg),  $n = 1\,050$  (Bayern) und  $n = 299$  (Nordrhein-Westfalen).

*Itemgewinnung.* Die Generierung der Itemmenge erfolgte in mehreren Phasen. Zunächst wurden Itemstämme formuliert, die den theoretischen Definitionen der zu erfassenden Merkmale entsprachen und diese möglichst distinkt indizieren sollten. Um Verständlichkeit für unsere Zielpopulation zu gewährleisten, bemühten wir uns, diese Stämme möglichst kurz, syntaktisch einfach und inhaltlich eindeutig zu fassen. Auf negativ formulierte Aussagen und komplexere Satzkonstruktionen wurde z. B. verzichtet. Diese Version wurde mit Hilfe „kognitiver Interviews“ (Prüfer & Rexroth, 2005) bei neun Kindern im Alter von 7 bis 15 Jahren geprüft. Hierbei wurden die Items den Kindern einzeln vorgelegt. Sie sollten für jedes Item mitteilen, ob und wie die Aussage verstanden wurde, wie sie antworten würden, ob sie andere Wörter benutzen würden und ob ihnen ein Beispiel zu der Aussage einfiele. Zur Auswertung wurden die Interviews auf Tonband aufgezeichnet. Auf der Basis der Interviews wurden einige Items umformuliert, weitere Items wurden neu aufgenommen. Mit dem Antwortformat, zu dem am Ende des Interviews ebenfalls Fragen gestellt wurden, kamen auch jüngere Kinder gut zurecht. Diese „Alpha-Version“ des MAI-KJ wurde in einer zweiten Phase mit insgesamt 884 Kindern diverser Schultypen und Altersgruppen getestet und auf der Grundlage der Rückmeldungen der betreuenden Personen kontinuierlich weiterentwickelt. Die hier gesammelten Erfahrungen gingen in die erste Version des MAI-KJ ein, die in unveränderter Form einer größeren Stichprobe (Stichprobe 1) vorgelegt wurde.

*Behandlung fehlender Werte.* Bei den nach dem Fallausschluss (siehe Abschnitt *Stichprobe*) verbleibenden Datensätzen wurden fehlende Werte mittels Predictive Mean Matching (PMM) imputiert. Die Imputation wurde mit MICE (van Buuren & Groothuis-Oudshoorn, 2011) vorgenommen. PMM ist für Items mit ihrem begrenzten Bereich diskreter Werte attraktiv, da bei diesem Verfahren die Imputationen auf tatsächlich beobachtete Werte limitiert werden. Da fehlende Werte nur vereinzelt auftraten, Items aufgrund ihrer hohen

Interkorrelationen sehr geeignete Ziele für Imputationsmethoden darstellen und die Stichproben sehr groß sind, lieferten multiple Imputationen praktisch jeweils die gleichen Mittelwerte, Varianzen und Kovarianzen unter den Items. Wir konnten uns daher auf die Analyse *einer* Imputation beschränken anstatt die Ergebnisse mehrerer Imputationen zu aggregieren.

### Literaturverzeichnis

- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67 (1), 1–48. doi:[10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)
- Prüfer, P. & Rexroth, M. (2005). *Kognitive Interviews*. Mannheim: Zentrum für Umfragen, Methoden und Analysen - ZUMA. URL: <http://www.ssoar.info/ssoar/handle/document/20147>
- van Buuren, S. & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45 (3), 1–67. doi:10.18637/jss.v045.i03

*Elektronisches Supplement 3. Explorative Faktorenanalysen des MAI-KJ*

Hinweise auf die Anzahl zu extrahierender Faktoren wurden mit Parallelanalysen gewonnen, die (wie die Faktorenanalysen) auf polychorischen Korrelationen aufbauten. Für die Faktorisierung wurde die Maximum-Likelihood-Methode eingesetzt, die bei größeren Stichproben genauere Kennwerte zu liefern verspricht als gängige Alternativen, wie z. B. die Hauptachsen- oder die Minimum-Residuals-Methode. Da besonders für Besorgnis, Aufgeregtheit und Traurigkeit von vornherein deutliche Korrelationen zu erwarten waren, wurden die Faktoren schiefwinklig rotiert. Auch zwischen Freude bzw. Anhedonie und den anderen Faktoren sollten substantielle Korrelationen bestehen, die jedoch für die Angstkomponenten niedriger ausfallen sollten als für die Traurigkeit (siehe Renner et al. 2018). Als Rotationsmethode wurde Promax mit dem (üblicherweise verwendeten) Exponenten von  $m = 4$  benutzt.

Parallelanalysen wurden sowohl nach dem Faktoren- als auch nach dem Hauptkomponentenmodell berechnet. Wie neuere Simulationsstudien zeigen, sind Hauptkomponentenanalysen (sofern sie auf polychorischen Korrelationen basieren) für diesen Zweck gut geeignet (Garrido, Abad & Ponsoda, 2013), obgleich für weitergehende Item- und Testanalysen „echte“ Faktorenmodelle verwendet werden sollten, da bei der Hauptkomponentenanalyse im Unterschied zur Faktorenanalyse die itemspezifischen Fehleranteile nicht berücksichtigt werden. Die Werte für die jeweils 20 Simulationsläufe der Parallelanalysen wurden zufällig aus den beobachteten Werten gezogen (Resampling), was konzeptuell angemessener ist als die Ziehung aus Normalverteilungen (bei größeren Stichproben sind die Unterschiede zwischen beiden Methoden allerdings marginal). Faktoren- und Parallelanalysen wurden mit dem R-Paket *psych* (Revelle, 2018) durchgeführt.

In Abbildung 1 sind die Ergebnisse der beiden Parallelanalysen einander gegenübergestellt. Die gestrichelten und mit Dreiecken verbundenen Linien zeigen die Eigenwertverläufe der ersten Faktoren bzw. Komponenten aus den Parallelanalysen. Da in der Faktorenanalyse nur die gemeinsame – nicht wie in der Hauptkomponentenanalyse die gesamte – Varianz betrachtet wird, sind empirische und simulierte Eigenwerte bei der Faktorenanalyse niedriger als bei der Hauptkomponentenanalyse. Aus dem gleichen Grund können Eigenwerte in der Faktorenanalyse auch geringfügig unter Null liegen, nämlich dann, wenn die extrahierten Faktoren die gemeinsame Varianz bereits erschöpfen, was in den vorliegenden Daten nach der Extraktion von fünf Faktoren der Fall ist. Die faktorenanalytisch berechneten Eigenwerte (durchgezogene Linien, Punkte) wiesen eindeutig auf eine Vier-Faktorenlösung hin, während die Eigenwerte der Hauptkomponentenanalyse auch mit drei Faktoren kompatibel waren. Aufgrund der Stichprobengröße überrascht es nicht, dass die Ergebnisse der Parallelanalysen recht genau mit dem „Kaiserkriterium“ übereinstimmten, nach dem Faktoren zu extrahieren sind, deren Eigenwerte über dem Eigenwert unkorrelierter Variablen liegt, was bei der Hauptkomponentenanalyse zu einem Grenzwert von ca. 1 und bei der Faktorenanalyse zu einem Grenzwert von ca. 0 führt.

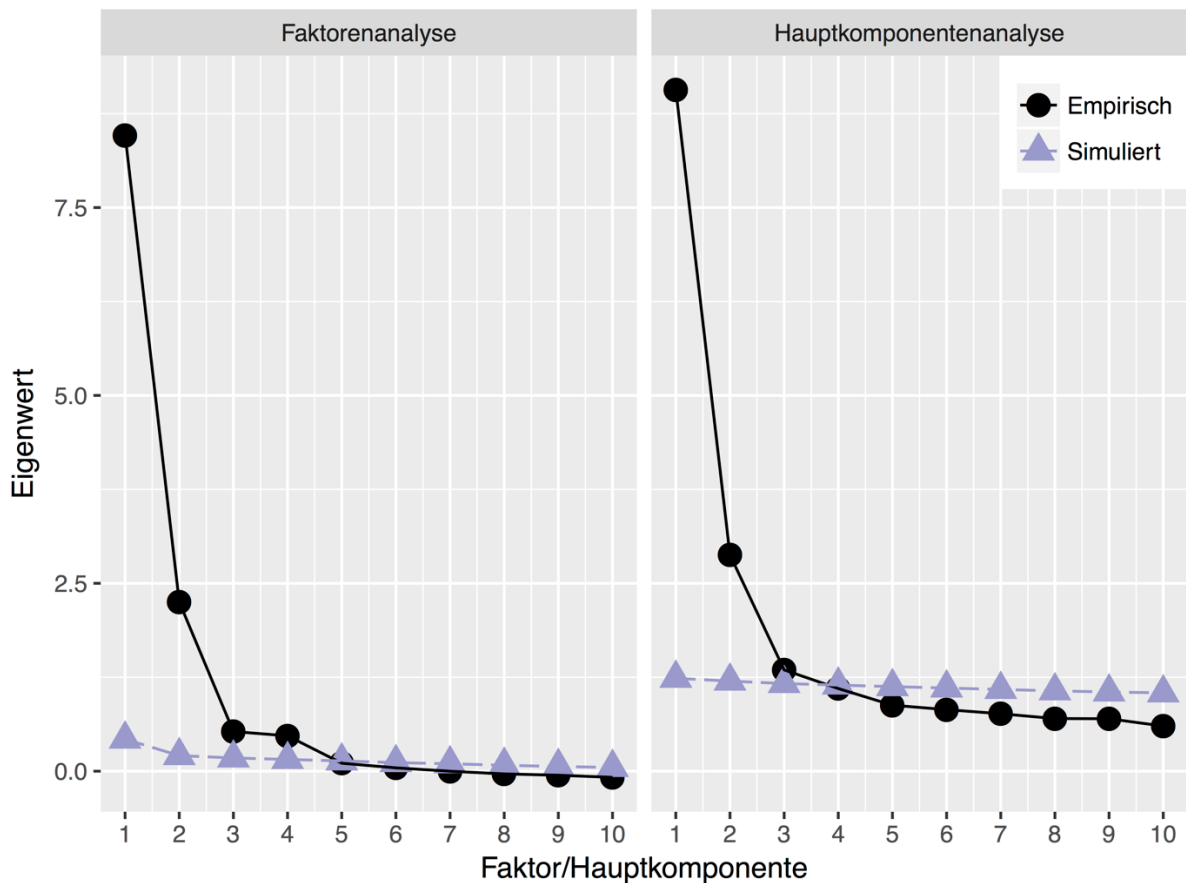


Abbildung 1. Parallelanalysen des MAI-KJ.

Wir entschieden uns für die Vier-Faktorenlösung, die den theoretischen Annahmen genau entsprach, da die Items eindeutig einem Besorgnis-, Aufregtheits-, Freude- oder Traurigkeitsfaktor zugeordnet wurden. Die alternative Drei-Faktorenlösung war dagegen psychologisch nicht sinnvoll interpretierbar; bei ihr fiel der Aufregtheitsfaktor heraus, die Markieritems dieses Faktors verteilten sich auf Besorgnis und Traurigkeit oder wiesen Mehrfachladungen auf beiden Faktoren auf. Tabelle 1 führt die Faktorladungen, Komplexitäten und Kommunalitäten der Items sowie die Eigenwerte und Korrelationen der vier Faktoren auf. Die Items sind hier für jeden Faktor nach Ladungshöhe sortiert. Wie ersichtlich ist, wurden alle Items den postulierten Faktoren in der erwarteten Weise zugeordnet. Von einer Ausnahme (Item 36) abgesehen, waren die Hauptladungen der Items größer als .40, die Nebenladungen relativ gering. Die Komplexitäten der Items fielen durchweg geringer aus als 2.

Tabelle 1. Explorative Faktorenanalyse (Promax-Rotation) des MAI-KJ (Allgemeiner Teil)

ID	$a_1$	$a_2$	$a_3$	$a_4$	$h^2$	$K$	Item
	FR	BE	AU	TR			
23	<b>.86</b>	.03	.00	.03	.69	1.01	Ich bin glücklich.
45	<b>.83</b>	-.04	.04	-.03	.72	1.01	Ich bin gut drauf.
55	<b>.83</b>	-.07	.10	.01	.66	1.04	Ich freue mich.
35	<b>.81</b>	-.06	.02	-.02	.70	1.01	Ich bin gut gelaunt.
1	<b>.75</b>	.03	.01	.03	.51	1.01	Ich bin fröhlich.
17	<b>.75</b>	.07	-.12	-.07	.68	1.09	Mir geht es gut.
27	<b>.74</b>	.01	.00	-.03	.57	1.00	Ich finde das Leben schön.
42	.05	<b>.70</b>	-.05	.15	.55	1.11	Ich mache mir Sorgen, dass andere schlecht über mich denken.
32	-.07	<b>.68</b>	.03	.00	.53	1.03	Ich mache mir Sorgen mich zu blamieren.
54	-.01	<b>.64</b>	.05	-.01	.45	1.01	Wenn mir etwas Unangenehmes passiert ist, gehen mir die Gedanken daran nicht aus dem Kopf.
46	-.03	<b>.62</b>	-.04	.06	.42	1.03	Ich bin besorgt, dass andere besser sind als ich.
29	.03	<b>.61</b>	.11	-.08	.40	1.11	Wenn mir etwas nicht gelungen ist, muss ich immer wieder daran denken.
24	-.05	<b>.50</b>	.25	-.02	.48	1.50	Ich habe Angst etwas falsch zu machen.
36	.03	<b>.39</b>	.10	.04	.22	1.16	Ich beschäftige mich damit, was noch passieren könnte.
12	.12	-.05	<b>.65</b>	.09	.39	1.13	Mein Herz klopft stark.
3	-.08	-.02	<b>.59</b>	-.13	.29	1.14	Ich bin nervös.
43	.02	.19	<b>.55</b>	-.14	.37	1.40	Meine Hände oder Beine zittern vor Aufregung.
26	.03	.12	<b>.51</b>	.17	.48	1.35	Ich habe ein unangenehmes Gefühl im Bauch.
4	-.10	-.04	<b>.45</b>	-.04	.20	1.14	Ich fühle mich angespannt.
6	-.06	-.07	<b>.40</b>	.04	.17	1.11	Ich bin unruhig.
39	-.08	.13	-.03	<b>.69</b>	.65	1.11	Ich bin einsam.
31	.00	.26	-.07	<b>.65</b>	.61	1.34	Ich fühle mich allein, auch wenn ich mit anderen zusammen bin.
28	-.07	-.04	.21	<b>.60</b>	.59	1.28	Ich bin traurig.
16	.11	-.02	.34	<b>.54</b>	.50	1.80	Ich habe das Gefühl, dass ich weinen möchte.
40	-.23	-.01	.06	<b>.52</b>	.52	1.41	Ich bin unglücklich.
Eigenwert	4.69	3.04	2.25	2.37			
Varianz	.19	.12	.09	.09			
Faktorkorrelationen							
		BE	AU	TR			
FR		-.40	-.42	-.64			
BE			.65	.58			
AU				.57			

Anmerkungen: Die Items sind nach Ladungshöhe sortiert. Die Hauptladungen von Markieritems sind fett gedruckt. ID = Itemkennung,  $a$  = Faktorladung,  $h^2$  = Kommunalität,  $K$  = faktorielle Komplexität, F = Faktor.  $N = 2\,594$ . Faktoreninterpretation: FR = Freude, BE = Besorgnis, AU = Aufgeregtheit, TR = Traurigkeit.

Ebenfalls erwartungsgemäß waren die Faktoren deutlich assoziiert, wobei die Korrelationen besonders für die Skalen mit „negativen“ Inhalten (Aufgeregtheit, Besorgnis, Traurigkeit) recht hoch ausfielen ( $r$ s zwischen .57 und .65). Der Zusammenhang der beiden Angstkomponenten Aufgeregtheit und Besorgnis war dabei mit  $r = .65$  am höchsten. Freude war am stärksten mit Traurigkeit korreliert,  $r = -.64$ , die Assoziationen der Freude mit Besorgnis und Aufgeregtheit sanken demgegenüber deutlich ab,  $r = -.40$  bzw.  $-.42$ . Damit erscheint auch die Bildung zweier übergeordneter Faktoren möglich, die Angst (Besorgnis, Aufgeregtheit) und Depression (Traurigkeit, mangelnde Freude bzw. Anhedonie) beinhalten.

### Literaturverzeichnis

- Garrido, L. E., Abad, F. J. & Ponsoda, V. (2013). A new look at Horn's parallel analysis with ordinal variables. *Psychological Methods*, 18, 454–474. doi:10.1037/a0030005
- Renner, K.-H., Hock, M., Bergner-Köther, R. & Laux, L. (2018). Differentiating anxiety and depression: The State-Trait Anxiety-Depression Inventory. *Cognition and Emotion*, 32(7), 1409–1423. doi:10.1080/02699931.2016.1266306
- Revelle, W. (2018) *Psych: Procedures for personality and psychological research*. R package version 1.8.12. URL <https://CRAN.R-project.org/package=psych>

*Elektronisches Supplement 4. Messinvarianz des MAI-KJ*

Wir berechneten vier Modelle, mit denen sogenannte konfigurale, schwache, starke und strikte Invarianz geprüft wurden (Cheung & Rensvold, 2002). Im ersten Modell, das als Referenzmodell für die in den weiteren Schritten untersuchten Modelle diente, konnten Faktorladungen und Schwellenparameter (bis auf die zur Identifikation der Skalen der Faktoren notwendigen Fixierungen jeweils einer Ladung und einer Schwelle pro Faktor auf 1) über die Gruppen frei variieren. Dieses Modell prüft *konfigurale Invarianz*, welche die Zahl der Faktoren und die Zuordnung der Items zu den Faktoren betrifft. Um die Identifizierbarkeit des Modells zu gewährleisten, wurden zusätzlich die Residualvarianzen aller Items fixiert (Muthén & Muthén, 1998–2012).

Im zweiten Modell wurden die Faktorladungen jedes Items über die Gruppen gleichgesetzt. Eine gute Passung dieses Modells indiziert *schwache Invarianz*, in der die Zusammenhänge zwischen den Items (bzw. der ihnen zugrunde liegenden kontinuierlichen Variablen) und den Faktoren für alle Items identisch sind. Bei Vorliegen schwacher Invarianz manifestieren sich Faktoren damit bei allen Gruppen in der gleichen Art und Weise. Auch in diesem Modell wurden, um Identifizierbarkeit sicherzustellen, die Residualvarianzen über die Gruppen fixiert.

Im dritten Modell wurden Faktorladungen und Schwellenparameter über die Gruppen gleichgesetzt und damit *starke Invarianz* geprüft. Die Gleichsetzung der Schwellenparameter impliziert die Annahme, dass die Gruppen die Antwortskala der Items in derselben Weise nutzen, also z. B. der Übergang von der ersten zur zweiten Antwortkategorie bei einem Item in allen Gruppen bei der gleichen Ausprägung auf dem Merkmalskontinuum stattfindet. Aufgrund der Restriktionen bei den Ladungen und Schwellen ist eine Fixierung der Residualvarianzen für die Identifizierbarkeit des Modells nicht notwendig. Diese wurden hier (bis auf die Fixierung auf 1 in einer Gruppe) frei gelassen.

Im vierten Modell wurde schließlich *strikte Invarianz* getestet. Hier werden Ladungen, Schwellen und Residualvarianzen der Items über die Gruppen gleichgesetzt. Dieses Modell unterscheidet sich vom Modell zur Prüfung der starken Invarianz durch die Fixierung der Residualvarianzen und prüft damit die zusätzliche Annahme, dass die Items in allen Gruppen mit der gleichen Genauigkeit messen.

Fitindices für diese Modelle sind in Tabelle 1 zusammengestellt.

*Tabelle 1.* Passung verschiedener Modelle zur Messinvarianz (WLSMV-Schätzung) bezüglich des Geschlechts und der Klassenstufe

Invarianz	$\chi^2$	<i>df</i>	CFI	$\Delta$ CFI	RMSEA	$\Delta$ RMSEA
Geschlecht						
Konfigural	4 380.06	538	.966		.044	
Schwach	3 978.75	559	.970	.004	.041	-.003
Stark	5 069.55	605	.961	-.009	.045	.004
Strikt	5 059.25	630	.961	.000	.044	-.001
Altersgruppe (Klassenstufe)						
Konfigural	5 129.34	807	.963		.047	
Schwach	4 713.45	849	.967	.004	.043	-.004
Stark	5 875.46	941	.958	-.009	.046	.003
Strikt	6 442.80	991	.954	-.004	.047	.001

*Anmerkungen:* Für die Bildung von Altersgruppen wurden die Klassenstufen 3 und 4 ( $n = 2\,538$ ), 5 bis 7 ( $n = 2\,834$ ) sowie 8 bis 10 ( $n = 1\,967$ ) zusammengefasst.  $\Delta$ CFI und  $\Delta$ RMSEA = Differenzen zwischen zunehmend strikteren Modellen.

## Literaturverzeichnis

- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. doi:10.1207/S15328007SEM0902\_5
- Muthén, L. K. & Muthén, B. O. (1998–2012). *Mplus user's guide*. (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Muthén, B. O. & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 25, 267–316. doi:10.2307/271070

Elektronisches Supplement 5. Korrelationen des MAI-KJ mit anderen Verfahren

Die Korrelationen der MAI-KJ-Skalen mit anderen Verfahren zur Erfassung angst- und depressionsbezogener Merkmale sind in Tabelle 1 zusammengestellt.

Da bei unseren Erhebungen Schulklassen untersucht wurden, überlagern sich in den Daten individuelle Differenzen und Gruppenunterschiede. Diese Überlagerung kann zu Verzerrungen in Korrelationen und anderen Assoziationsmaßen führen, wenn die Verteilungen der Variablen zwischen den Gruppen sehr heterogen sind, die Gruppen also z. B. hinsichtlich ihrer Mittelwerte stark differieren. Wir prüften daher für jede Variable (Item- und Skalenscores), ob die Heterogenität zwischen den Schulklassen so groß ist, dass sie in den Analysen berücksichtigt werden sollte. Hierfür schätzten wir Designeffekte auf der Basis der Intraklassenkorrelationen (ICC) und der mittleren Gruppengröße ( $n$ ) nach der Näherungsformel  $1 + ICC(n - 1)$ . Designeffekte repräsentieren das Ausmaß, in dem die Varianz einer Variablen durch die Gruppenerhebung gegenüber der Varianz einer einfachen Zufallsstichprobe gleichen Umfangs (für welche der Designeffekt 1 wäre) erhöht ist. Designeffekte größer als 2 gelten dabei als kritisch (Maas & Hox, 2005; Muthen & Satorra, 1995). Für die Items und Skalen des MAI-KJ erreichte kein Designeffekt die kritische Grenze. Bei den zur Validierung herangezogenen Skalen ergaben sich dagegen einige erhöhte Werte. Wir berechneten deshalb der Einfachheit halber generell Within-Group-Korrelationen zwischen den MAI-KJ-Skalen und den anderen Variablen, indem wir alle Variablen vor der Berechnung der Korrelation gruppenspezifisch zentrierten, was Mittelwertsdifferenzen zwischen den Gruppen, welche eventuell zu einer Überschätzung der Zusammenhänge führen können, eliminiert. Diese Korrelationen waren allerdings nahezu identisch mit den Korrelationen der Ausgangsvariablen.

Tabelle 1. Korrelationen der MAI-KJ-Skalen mit anderen Skalen zur Erfassung angst- und depressionsbezogener Merkmale

Variable	AN	AU	BE	DE	TR	FR	$t^a$
AFS ( $n = 433, k = 23$ )							
Prüfungsangst	.52*	.42*	.49*	.31*	.34*	-.23*	5.52*
Manifeste Angst	.64*	.50*	.62*	.48*	.54*	-.34*	4.67*
Schulunlust	.24*	.23*	.20*	.40*	.26*	-.39*	3.84*
Soziale Erwünschtheit	-.11*	-.08	-.11*	-.24*	-.10*	.27*	3.05*
KAT-II ( $n = 325, k = 17$ )							
Gesamtscore	.71*	.61*	.66*	.59*	.62*	-.46*	3.32*
SASC-R-D ( $n = 530, k = 28$ )							
Fear of Negative Evaluation	.58*	.38*	.60*	.45*	.51*	-.31*	3.71*
Social Avoidance and Distress	.47*	.35*	.46*	.36*	.41*	-.25*	2.91*
DTK-II ( $n = 486, k = 36$ )							
Dysphorie/Selbstwertprobleme	.56*	.37*	.59*	.58*	.63*	-.41*	0.42
Agitiertes Verhalten	.25*	.21*	.24*	.32*	.25*	-.30*	1.52
Müdigkeit/autonome Reaktionen	.41*	.34*	.39*	.38*	.38*	-.30*	0.72
DTGA ( $n = 109, k = 6$ )							
Gesamtscore Depressivität	.53*	.48*	.48*	.58*	.59*	-.49*	0.72

Anmerkungen: AN = Angst, AU = Aufgeregtheit, BE = Besorgnis, DE = Depression, TR = Traurigkeit, FR = Freude. AFS = Angstfragebogen für Schüler, KAT-II = Kinder-Angst-Test-II, SASC-R-D = Social Anxiety Scale for Children – revised, DTK-II = Depressionstest für Kinder, DTGA = Depressionstest für Kinder im Grundschulalter.

<sup>a</sup>  $t$ -Wert zur Prüfung der Korrelationsdifferenz der betreffenden Skala mit Angst (AN) und Depression (DE).

\* $p < .05$  (zweiseitige Testung).

Literaturverzeichnis

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology, 1*(3), 86–92. <https://doi.org/10.1027/1614-1881.1.3.86>

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25*, 267–316. <https://doi.org/10.2307/271070>

