

Fragebogen zur Erfassung des Humors von Lehrkräften aus Schülersicht (HUMLAS): Konstruktion und Validierung

Sonja Bieg und Markus Dresel

Zusammenfassung. Eine deutschsprachige Adaption der „Teacher Humor Scale“ (Frymier, Wanzer & Wojtaszczyk, 2008) erfasst mit 17 Items vier Formen des schülerperzipierten Lehrerhumors: lerngegenstandsbezogener Humor, Humor ohne Bezug zum Lerngegenstand, selbstabwertender Humor und aggressiver Humor. Eine Studie mit 890 Gymnasiasten der 9. und 10. Jahrgangsstufe aus 41 Schulklassen erbrachte klare Hinweise auf die Messgüte des Verfahrens. Konfirmatorische Faktorenanalysen auf Schülerebene bestätigten die vierfaktorielle Struktur. Ergebnisse von explorativen Zwei-Ebenen-Faktorenanalysen standen auch auf Klassenebene im Einklang mit der angenommenen Struktur des Lehrerhumors. Die auf dieser Basis gebildeten Subskalen verfügten über eine gute Reliabilität. Die Analyse der Zusammenhänge mit Schülermerkmalen und anderen Lehrer-/Unterrichtsmerkmalen erbrachte Hinweise auf die konvergente und die diskriminante Validität des Verfahrens. Mit dem „Fragebogen zur Erfassung des Humors von Lehrkräften aus Schülersicht“ (HUMLAS) liegt damit ein reliables, valides und ökonomisches Instrument zu einem Aspekt von Unterricht vor, zu dem es bislang kein differenziertes Verfahren gab.

Schlüsselwörter: Lehrer, Humor, Humorformen, Schülerwahrnehmung, Fragebogen

Construction and Validation of the German Questionnaire to Assess Students' Perceptions of Teacher Humor

Abstract. The construction and validation of the German Questionnaire to Assess Student Perceptions of Teacher Humor, a translation and adaptation of the Teacher Humor Scale (Frymier, Wanzer & Wojtaszczyk, 2008), is reported. The scale consists of 17 items assessing four types of teacher humor, namely humor related to course material, humor unrelated to course material, self-disparaging humor, and aggressive humor. A study with 890 high school students in Years 9 and 10 from 41 different school classes was conducted to test the quality of the measuring instrument. Results from confirmatory factor analyses on the student level and two-level exploratory factor analyses supported the hypothesized structure with four factors. Subsequently constructed subscales demonstrated a good internal consistency. Convergent and discriminant validity was supported by correlations with student, teacher and class characteristics. It can be concluded that the new questionnaire is a reliable, valid and parsimonious measuring instrument, which can be used in several contexts.

Keywords: teacher, types of humor, student perceptions, questionnaire

Humor spielt im Unterricht eine nicht zu unterschätzende Rolle (z.B. Banas, Dunbar, Rodriguez & Liu, 2011). Bringen Lehrkräfte Witz und Heiterkeit in den Unterricht ein, bietet ihnen dies eine Möglichkeit, soziale, motivationale und kognitive Aspekte günstig zu beeinflussen und ihre Lehrziele besser zu erreichen (vgl. Dresel et al., 2014; Ziv, 1979). Nach Frymier, Wanzer und Wojtaszczyk (2008) muss zwischen verschiedenen Humorformen bei Lehrpersonen differenziert werden – demnach ist Lehrerhumor ein multidimensionales Konzept, das sowohl angemessene Humorformen (z.B. witzige Beispiele für Unterrichtsinhalte) wie auch unangemessene Humorformen (z.B. Lernende lächerlich machen) umfasst (vgl. Martin, Puhlik-Doris, Larsen, Gray & Weir, 2003; Wanzer, Frymier, Wojtaszczyk & Smith, 2006; Wanzer, Frymier & Irwin, 2010). Bislang gibt es jedoch kein deutschsprachiges Instrument,

mit dem eine differenzierte Erfassung des Lehrerhumors valide möglich ist. Die wenigen Studien zum Lehrerhumor im deutschsprachigen Raum (z.B. Rißland, 2002) nutzten fast ausschließlich Lehrerselbstberichte. Eine Ausnahme stellt lediglich eine Studie von Dresel et al. (2014) dar, für die der Humor von Lehrkräften mit Hilfe der Wahrnehmungen von Schülerinnen und Schülern operationalisiert wurde (allerdings nicht nach verschiedenen Formen differenziert). Ein Messzugang mit Hilfe von Schülerperceptionen bietet sich zur Überwindung der Verzerrungen an, die häufig mit Lehrerselbstberichten verbunden sind (vgl. Helmke, 2009). Zudem bietet dies den Vorteil, dass damit jener Anteil des Lehrerhumors erfasst wird, der für das Erleben und Lernverhalten von Schülerinnen und Schülern vorrangig relevant ist. Ziel der vorliegenden Arbeit ist vor diesem Hintergrund die Entwicklung und

| Humorfunktionen und -formen in Alltagskontexten nach Martin et al. (2003) | | | Lehrerhumor nach Frymier et al. (2008) | | Fragebogen zur Erfassung des Humors von Lehrkräften aus Schülersicht |
|--|----------------------------------|----------------------------|--|--|--|
| Funktion | Verträglichkeit | Form | Bezug Lern- gegenstand | Form | |
| Verbesserung soziale Beziehungen | verträglich | Affiliativer Humor | mit Bezug | Humor mit Unterrichtsbezug | Lerngegenstands- bezogener Humor |
| | | | ohne Bezug | Humor ohne Unterrichtsbezug | Humor ohne Bezug zum Lerngegenstand |
| Verbesserung soziale Beziehungen | auf Kosten der eigenen Person | Selbstabwertender Humor | ohne Bezug | Humor, der eigene Person herabsetzt | Selbstabwertender Humor |
| Verbesserung Selbst | verträglich | Selbstgerichteter Humor | nicht berücksichtigt (kaum relevant für sozialen Kontext) | | |
| Verbesserung Selbst | auf Kosten anderer Personen | Aggressiver Humor | ohne Bezug | Humor, der andere Personen herabsetzt | Aggressiver Humor |
| | | | ohne Bezug | Offensiver Humor | |

Abbildung 1. Übersicht über Humorformen im Modell von Martin et al. (2003) für Alltagskontexte, im Instrument zur Erfassung des Lehrerhumors von Frymier et al. (2008) sowie in der theoretischen Konzeption des „Fragebogens zur Erfassung des Humors von Lehrkräften aus Schülersicht“ (HUMLAS).

Validierung eines Fragebogenverfahrens, mit dem relevante Formen des Lehrerhumors differenziert aus Schülersicht erfasst werden. Basierend auf den Modellvorstellungen von Martin et al. (2003) sowie den im Kontext US-amerikanischer Colleges entwickelten Skalen von Frymier et al. (2008) wurde dazu ein deutsches Messinstrument entwickelt, das anhand einer Studie im Deutschunterricht der 9. und 10. Jahrgangsstufe empirisch überprüft wurde.

Definition, Funktionen und Formen von Humor

Nach Martin (2007) umfasst Humor die Kommunikation von inkongruenten Bedeutungen, die in irgendeiner Weise amüsierend sind, und basiert rezipientenseitig auf der Wahrnehmung und Auflösung dieser Inkongruenzen (vgl. Berlyne, 1960; LaFave, Haddad & Maesen, 1996). Für soziale Kontexte im Allgemeinen findet sich in der Literatur eine große Zahl an Forschungsarbeiten zur Nutzung von Humor und dessen Wirkungen für Humorproduzenten und -rezipienten (Überblicke bei Lefcourt, 2001; Martin, 2007). Für die vorliegende Arbeit sind zwei theoretische Aspekte von besonderer Bedeutung (Martin, 2003): (1) Humor sollte als multidimensionales Konzept mit verschiedenen Humorformen aufgefasst werden. (2) Zur Erklärung der Wirkungen von Humor für Produzenten und Rezipienten ist eine funktionale Perspektive, die zwischen verschiedenen Intentionen und Anlässen für die Humorproduktion differenziert, geeigneter als eine Perspektive, die nach verschiedenen Humoringhalten (z.B. Witze, Anekdoten, Kommentare etc.) unterscheidet.

Das theoretische Verständnis des Lehrerhumors in dieser Arbeit geht auf das 2 x 2-Modell von Martin et al. (2003) zum Humor in Alltagskontexten zurück (Abbildung 1).

Martin und Kollegen unterscheiden dabei, ob Humor genutzt wird, um soziale Beziehungen zu fördern (soziale Funktion) oder um das Selbst zu fördern (intrapyschische Funktion). Auf einer zweiten Dimension unterscheiden sie, ob Humor für alle Beteiligten verträglich ist oder nicht (d.h. entweder auf Kosten anderer Personen oder des Humorproduzenten selbst geht). Aus diesen beiden Dimensionen leiteten Martin et al. (2003) vier verschiedene Formen von Humor ab: Dies sind zunächst „affiliativer Humor“ (Martin et al., 2003, S. 53) (z.B. Erzählen von freundlichen Witzen) und „selbstabwertender Humor“ (Martin et al., 2003, S. 54) (z.B. Selbstironie; Anekdoten über eigene Missgeschicke), die beide primär dazu genutzt werden, die Beziehungen mit Interaktionspartnern zu fördern. Während die erste Humorform verträglich für alle Beteiligten ist, geht die zweite auf Kosten des Selbst und hat auch empirisch nachweisbar negative Wirkungen für den Humorproduzenten (vgl. Kubie, 1971). Weiterhin ist „selbstgerichteter Humor“ (Martin et al., 2003, S. 53) (Betrachtung des Lebens aus einer generell humorvollen Perspektive) zu nennen, der primär zur Stressbewältigung und zum Selbstwertschutz genutzt wird; charakteristisch für ihn sind vor allem bestimmte intrapsychische Prozesse (z.B. humorvolle Gedanken), weshalb er für andere Personen schwer wahrnehmbar ist (Kuiper, Martin & Olinger, 1993). Schließlich zielt „aggressiver Humor“ (Martin et al., 2003, S. 54) (z.B. andere Personen lächerlich ma-

chen) auf die Aufwertung des Selbst auf Kosten Anderer; er hat negative Wirkungen für soziale Beziehungen und die Personen, auf die der Humor gerichtet ist (Martin et al., 2003; Zillman, 1983).

Konzeption und Erfassung verschiedener Formen des Lehrerhumors

Zum Humor von Lehrpersonen liegt mittlerweile eine durchaus nennenswerte Zahl an Forschungsarbeiten vor, die allerdings überwiegend an US-amerikanischen Colleges oder Universitäten durchgeführt wurden, meist ein unidimensionales Verständnis des Lehrerhumors anlegten und sich dabei häufig primär dessen Inhalten zuwandten (für Überblicke siehe Banas et al., 2011; Booth-Butterfield & Wanzer, 2010).

Eine Ausnahme von der unidimensionalen Perspektive stellen die Forschungen der Gruppe um Wanzer und Frymier dar (z. B. Frymier et al., 2008; Wanzer & Frymier, 1999; Wanzer et al., 2006; Wanzer et al., 2010), die zentral für die vorliegende Arbeit sind – obwohl sie ebenfalls im Hochschulforschungsbereich angesiedelt waren. Anhand einer qualitativen Befragung von Studierenden schlossen Wanzer et al. (2006) auf verschiedene Humorformen, die in der universitären Lehre eingesetzt werden, sowie die Angemessenheit dieser Humorformen. Dabei zeigte sich, dass affiliativer Humor und selbstabwertender Humor der Lehrperson von Studierenden als angemessen erlebt werden, etwa weil dies das Klima in der Gruppe verbessert. Dies galt insbesondere, wenn ein Bezug zum Lerngegenstand bestand (z. B. wenn humorvolle Erklärungen für Lerninhalte präsentiert wurden); dies trug aus Studierenden-sicht dazu bei, die Interessanztheit des Lerninhalts zu steigern sowie Informationsenkodierung und -abruf zu fördern. In einer quantitativen Nachfolgestudie zeigte sich, dass affiliativer Lehrerhumor, der auf den Lerngegenstand bezogen ist, positiv mit dem Lernzuwachs der Lernenden assoziiert ist (Wanzer et al., 2010). Aggressiver Lehrerhumor auf Kosten Anderer (anwesende oder nicht anwesende Personen, Personengruppen) wurde von Studierenden dagegen als unangemessen wahrgenommen und korrelierte nicht mit dem Lernzuwachs.

Basierend auf diesen Ergebnissen konstruierten Frymier et al. (2008) ein Fragebogenverfahren, mit dem Studierende den Humor von Lehrpersonen beurteilen können. Insbesondere berücksichtigten sie dabei neben der sozialen und der auf das Selbst gerichteten Funktion auch die auf die Vermittlung des Lerngegenstands gerichtete Funktion des Lehrerhumors, die spezifisch für instruktionale Kontexte ist (vgl. Abbildung 1). Im Anschluss an die Ergebnisse von Wanzer et al. (2006) differenzierten sie zwischen „Humor mit Unterrichtsbezug“ und „Humor ohne Unterrichtsbezug“ (Frymier et al., 2008, S. 275). Daneben operationalisierten sie Humor, mit dem sich die Lehrperson selbst

herabsetzt, und solchem, mit dem sie andere Personen herabsetzt. Selbstgerichteter Humor, der von außen schwer wahrnehmbar ist, wurde aufgrund seiner geringen Relevanz für Lehr-Lernprozesse nicht berücksichtigt. Frymier et al. (2008) untersuchten die Dimensionalität des von ihnen entwickelten Fragebogens (ausschließlich) anhand explorativer Methoden (explorative Faktorenanalyse der individuellen Wahrnehmungen der Studierenden unter Anwendung des Eigenwertkriteriums; keine Berücksichtigung der Lehrpersonen-/Kursebene). Die Ergebnisse standen im Einklang mit den vier angenommenen Formen des Lehrerhumors. Daneben fand sich aber auch noch ein fünfter, theoretisch nicht klar abgegrenzter Faktor, der sich u. a. auf die Nutzung vulgärer Sprache bezog und als besonders starke Form des aggressiven Humors verstanden werden kann („offensiver Humor“) (Frymier et al., 2008, S. 275). Das Instrument von Frymier et al. (2008) wies für alle Faktoren eine gute bis sehr gute Reliabilität auf (Cronbachs $\alpha = .80 - .93$).

Offen ist bislang die Übertragbarkeit der im US-amerikanischen College-Bereich gewonnenen Befunde auf den Kontext des schulischen Unterrichts und auf den deutschsprachigen Kulturraum. Wir erwarten, dass die vier aus einer theoretisch-funktionalen Perspektive auf Lehrerhumor abgeleiteten Faktoren von Frymier et al. (2008) auch im hiesigen Schulkontext voneinander abgrenzbare Humorformen darstellen und von Relevanz sind – auch wenn diese aufgrund des unterschiedlichen Alters der Lernenden und der Kulturabhängigkeit von Humor möglicherweise in anderen spezifischen Humorphänomenen zum Ausdruck kommen (Davies, 1998; Hill, 1988; Neuliep, 1991; Nevo, Nevo & Yin, 2001). Als übertragbar erachten wir auch die Annahme, dass selbstgerichteter Humor von Lehrpersonen wenig relevant im Kontext des Unterrichts ist. Beim offensiven Humor wichen wir jedoch von der Konzeption von Frymier und Kollegen ab: Hier ist bedeutsam, dass das höhere Alter von Studierenden und die andere Beziehungsstruktur in Lehrveranstaltungen im Hochschulforschungsbereich Dozierenden weniger Restriktionen bezüglich des Humorphänomens auferlegt als Lehrkräften an Schulen im Primar- oder Sekundarbereich (Hill, 1988). Damit korrespondierend erbrachte die Befragung von Neuliep (1991), dass Lehrkräfte an Schulen im Vergleich zu Dozierenden an Hochschulen kaum offensive Humorphänomenen einsetzten. Vor dem Hintergrund dieser kontextuellen Unterschiede und aufgrund der mangelnden theoretischen Fundierung des offensiven Humors sowie seiner schwachen empirischen Begründung erscheint es nicht gerechtfertigt, aggressiven Lehrerhumor in eine weniger offensive und eine offensivere Form zu differenzieren. Damit umfasst unser theoretisches Konzept vier relevante und von Schülerinnen und Schülern wahrnehmbare Formen des Lehrerhumors (vgl. Abbildung 1): *lerngegenstandsbezogener Humor*, *Humor ohne Bezug zum Lerngegenstand*, *selbstabwertender Humor* und *aggressiver Humor*.

Zur Erfassung der für das subjektive Erleben und (Lern-)Verhalten von Schülerinnen und Schülern primär bedeutsamen Aspekte des Lehrerhumors sowie zur Vermeidung der mit Lehrerselbstberichten verbundenen methodischen Probleme bietet sich die Nutzung von Fremdbeurteilungen durch Lernende an (vgl. Marsh et al., 2012). Dass das Instrument von Frymier et al. (2008) diesen Ansatz nutzt, ist eine seiner Stärken. Allerdings wurden in den Arbeiten von Frymier et al. (2008) und Wanzer et al. (2010) ausschließlich die individuellen Wahrnehmungen der Studierenden fokussiert und die vorrangig interessierende Ebene der Lehrperson sowie die geschachtelte Datenstruktur als solches ignoriert (mehrere Lernende beurteilen den Humor einer Lehrperson). Damit fehlt bislang eine Separierung der Effekte auf Ebene der Lehrpersonen (geteilte Wahrnehmung aller Lernenden) von jenen auf Ebene der Lernenden (von der geteilten Wahrnehmung abweichende Individualwahrnehmung, beeinflusst z.B. durch Erwartungseffekte). Darüber hinaus ist nichts über die Messgüte des Verfahrens von Frymier et al. (2008) auf der Ebene der Lehrpersonen bekannt.

Ziele und Hypothesen

Die Ergebnisse der bisherigen Forschung, die überwiegend im Hochschulbereich angesiedelt war, verweisen darauf, dass Lehrpersonen verschiedene Humorformen einsetzen, dass diese von Lernenden differenziert wahrgenommen werden können und sie für Unterricht und Lernverhalten relevant sind. Da allerdings bislang kein entsprechendes deutschsprachiges Verfahren vorliegt, war die Entwicklung und Validierung einer deutschen Adaption der „Teacher Humor Scale“ von Frymier et al. (2008) das Ziel der vorliegenden Arbeit, wobei dabei auch Schwächen der Originalskala überwunden werden sollten. Im Zusammenhang der Validierung prüften wir – ausgehend von den dargestellten Überlegungen und Befunden – drei Hypothesen.

Hypothese H1 bezieht sich auf die faktorielle Validität des Verfahrens und reflektiert das vorgestellte theoretische Verständnis des Lehrerhumors:

H1: Schülerperzipierter Lehrerhumor umfasst vier Humorformen: *lerngegenstandsbezogener Humor*, *Humor ohne Bezug zum Lerngegenstand*, *selbstabwertender Humor* und *aggressiver Humor*.

Hypothese H2 thematisiert die konvergente Validität des Verfahrens, zu deren Beurteilung theoretisch mit dem Lehrerhumor eng assoziierte Aspekte fokussiert wurden:

H2: Schülerwahrnehmungen des Lehrerhumors stehen im Zusammenhang mit Schülerwahrnehmungen des Lehrerenthusiasmus und der Interessantheit des Unterrichts (*lerngegenstandsbezogener Humor*, *Humor ohne Bezug zum Lerngegenstand* und *selbstabwertender Humor*: positive Zusammenhänge; *aggressiver Humor*: negative Zusammenhänge).

Der Enthusiasmus von Lehrkräften (im Sinne eines dispositionalen Lehrermerkmals, das sich darauf bezieht, wie viel Freude und Begeisterung Lehrkräfte während des Unterrichts erleben; vgl. Kunter et al., 2008) bezieht sich ebenso wie deren Humor auf ihr emotionales Erleben und hat überdies ähnliche Effekte auf sozio-emotionale und motivationale Dimensionen des Unterrichts (Dresel et al., 2014; Turner et al., 2002). Anzunehmen ist, dass Lehrerhumor darüber hinaus auch mit der Interessantheit des Unterrichts assoziiert ist, da die für Humor kennzeichnende Inkongruenz von Informationen eines der Aufgabencharakteristika ist, das situationales Interesse im Unterricht fördern kann (z.B. Berlyne, 1963; Hidi & Renninger, 2006). Diesen Überlegungen entsprechend korrelierte der unidimensional erfasste Lehrerhumor in der Studie von Dresel et al. (2014) moderat bis stark mit dem Lehrerenthusiasmus und der Interessantheit des Unterrichts.

Hypothese H3 adressiert schließlich die diskriminante Validität des Verfahrens, zu deren Beurteilung Unterrichts- und Schülermerkmale herangezogen wurden, die theoretisch nicht mit dem Lehrerhumor assoziiert sind:

H3: Schülerwahrnehmungen des Lehrerhumors stehen in keinem substanziellen Zusammenhang mit Schülerwahrnehmungen der Zeitnutzung im Unterricht sowie der Humororientierung und der Selbstwirksamkeit von Schülerinnen und Schülern.

Lehrkräfte können Humor dazu nutzen, bestimmte Lehrfunktionen des Unterrichts (Klauer, 1985) zu sichern (z.B. Motivationsfunktion); für bestimmte andere Lehrfunktionen ist aber nicht anzunehmen, dass er hilfreich ist (siehe Dresel et al., 2014, zur Frage, zu welchen Lehrfunktionen Humor beitragen kann). Letzteres gilt insbesondere für die übergeordnete Funktion der Steuerung und Kontrolle von Lernprozessen, die sich u. a. in der Zeitnutzung im Unterricht (*Time on Task*) manifestiert. Dieses theoretisch vom Lehrerhumor unabhängige Unterrichtsmerkmal fokussiert, wie viel Zeit im Unterricht auf den Lerngegenstand und wie viel Zeit für andere, für das eigentliche Unterrichtsthema irrelevante Aspekte verwendet wird (Helmke, 2009). Die Humororientierung und die Selbstwirksamkeit der Lernenden wurden betrachtet, da stets die Möglichkeit besteht, dass Fremdurteile (hier: des Lehrerhumors) von Merkmalen der Beurteiler (hier: der Lernenden) abhängen. Schülerinnen und Schüler mit ausgeprägtem Sinn für Humor könnten humorvolle Äußerungen ihrer Lehrkräfte entweder als besonders lustig wahrnehmen (Ähnlichkeitseffekt; könnte auch deshalb auftreten, weil diese Schülerinnen und Schüler darin geübt sind, Inkongruenzen aufzulösen) oder sie als weniger lustig wahrnehmen als es ihre Mitschülerinnen und Mitschüler tun (Kontrasteffekt; vgl. Booth-Butterfield & Booth-Butterfield, 1991). Ebenso könnte eine hohe Selbstwirksamkeit bedingen, dass die Lehrkraft, ihr Unterricht und somit auch ihr Humor generell positiver

wahrgenommen werden – möglicherweise auch aufgrund eines allgemeinen Sympathieeffektes, der beispielsweise aus positiven Leistungsbeurteilungen resultieren könnte, die auch das Wirksamkeitserleben der Schülerinnen und Schüler fördern (vgl. Helmke, 2009; Schwarzer & Jerusalem, 2002). Für die diskriminante Validität des Verfahrens spricht, wenn derartige Urteilerffekte hinreichend klein sind.

Methode

Konstruktion des Fragebogens

Die Items der „Teacher Humor Scale“ von Frymier et al. (2008) wurden ins Deutsche übersetzt und an den deutschen Kulturraum angepasst. Da es sich bei der Originalskala um ein Instrument zum Einsatz in US-amerikanischen Colleges handelt, war die Passung einiger Items zum deutschen Schulunterricht fraglich. Deshalb wurde der Itempool um einen um ein wenig inhaltsvalides Item aus der Skala „Humor ohne Unterrichtsbezug“ (Frymier et al., 2008, S. 275) gekürzt („Unser Lehrer verwendet kritischen, zynischen oder sarkastischen Humor zu allgemeinen Themen“) und zum Anderen um vier neu konstruierte Items ergänzt. Bei der Item-Neukonstruktion wurde darauf geachtet, dass sie den Spezifika des schulischen Unterrichts und dem jüngeren Alter der Lernenden gerecht werden. Um eine einfache Adaption an unterschiedliche Domänen (Schulfächer) zu ermöglichen, wurde die Fachspezifität im Itemstamm repräsentiert, die Items selbst aber allgemein formuliert (Beispiel: „Unsere Deutschlehrerin/Unser Deutschlehrer ... verwendet lustige Beispiele im Unterricht“). Die insgesamt 26 Items der Konstruktionsversion sind in Tabelle 1 aufgeführt. Analog zum Originalinstrument wurden die Items mit einer Häufigkeitsskala mit den Abstufungen 1 (*nie*), 2 (*selten*), 3 (*gelegentlich*), 4 (*oft*) und 5 (*sehr oft*) vorgegeben.

Stichprobe und Untersuchungsablauf

Zur Überprüfung der Messgüte der Skala wurde eine Untersuchung im Schulfach Deutsch durchgeführt. Die Stichprobe bestand aus 890 Schülerinnen und Schülern aus 41 Schulklassen der 9. und 10. Jahrgangsstufe an 11 Gymnasien in Baden-Württemberg, Bayern und Thüringen. Der Anteil der Mädchen lag bei 47 %. Das Durchschnittsalter der Schülerinnen und Schüler betrug 16.1 Jahre ($SD = 0.85$). Die Teilnahme an der Studie war freiwillig und für jeden Jugendlichen lag das schriftliche Einverständnis der Eltern vor (Teilnahmequote: 86 %). Die Untersuchung erfolgte im Rahmen einer regulären Unterrichtsstunde und wurde von geschulten Testleitern durchgeführt.

Messinstrumente zur Überprüfung der konvergenten und diskriminanten Validität

Enthusiasmus der Lehrkräfte. Der schülerperzipierte Enthusiasmus der Lehrkräfte für das Unterrichten wurde mit einem auf das Fach Deutsch adaptierten Instrument von Baumert et al. (2008) erfasst (Beispielitem: „Unser Lehrer ist selbst vom Fach Deutsch begeistert“). Die drei Items wurden mit einer fünfstufigen Likert-Skala vorgegeben, die von 1 (*stimme überhaupt nicht zu*) bis 5 (*stimme völlig zu*) reichten. Die interne Konsistenz war mit $\alpha = .85$ gut, ebenso lagen die übrigen Skalenkennwerte in einem angemessenen Bereich ($M = 3.91$; $SD = 0.88$; Intraklassenkorrelation $ICC1 = .20$).

Interessantheit des Unterrichts. Eine sechs Items umfassende Skala von Ditton und Merz (1999) erfasste, inwieweit die Lehrperson aus Sicht der Lernenden den Unterricht interessant gestaltet (Beispielitem: „Unser Lehrer gestaltet den Deutschunterricht abwechslungsreich“). Es kam eine fünfstufige Antwortskala zum Einsatz, die von 1 (*nie*) bis 5 (*immer*) reichte. Die Skala wies mit $\alpha = .91$ eine sehr gute interne Konsistenz auf ($M = 2.65$; $SD = 0.88$; $ICC1 = .34$).

Zeitnutzung im Unterricht. Der Umfang der für den Lerngegenstand genutzten Zeit im Unterricht wurde mit dem bereits in der Studie „Bildungsverläufe und psychosoziale Entwicklung im Jugendalter“ (BIJU) eingesetzten und negativ formulierten Einzelitem „Bei uns im Deutschunterricht wird viel Zeit vertrödelte“ erfasst (Baumert, Gruehn, Heyn, Köller & Schnabel, 1997). Verwendung fand eine Antwortskala, die von 1 (*stimmt gar nicht*) bis 6 (*stimmt völlig*) reichte. Die Werte wurden rekodiert, so dass hohe Werte eine gute Zeitnutzung repräsentierten. Die deskriptiven Kennwerte lagen im angemessenen Bereich ($M = 3.87$; $SD = 1.46$; $ICC1 = .22$).

Humororientierung der Schülerinnen und Schüler. Der Grad, zu dem sich die Schülerinnen und Schüler selbst als humorvoll verstehen und zu dem sie in der sozialen Interaktion Humor einsetzen, wurde mit einer Übersetzung der sechs Items umfassenden „Humor Orientation Scale“ von Booth-Butterfield und Booth-Butterfield (1991) erfasst. Ein Beispielitem lautet: „Ich erzähle regelmäßig Witze oder lustige Geschichten, wenn ich mit anderen zusammen bin“. Die Antworten wurden mit einer fünfstufigen Antwortskala von 1 (*stimmt gar nicht*) bis 5 (*stimmt völlig*) erfasst. Mit $\alpha = .85$ war die interne Konsistenz gut ($M = 3.29$, $SD = 0.75$; $ICC1 = .03$).

Selbstwirksamkeitsüberzeugung der Schülerinnen und Schüler. Die Selbstwirksamkeit von Schülerinnen und Schülern im Fach Deutsch wurde mit einer modifizierten Skala von Jerusalem und Satow (1999) erhoben. Sie umfasst sechs Items und erfasst das Ausmaß der Überzeugung, über die notwendigen personalen Ressourcen zur

Bewältigung schwieriger Aufgaben im Fach Deutsch zu verfügen (Beispielitem: „Ich kann auch die schwierigen Aufgaben in Deutsch lösen, wenn ich mich anstreng“). Alle Items waren auf einer vierstufigen Antwortskala von 1 (*trifft nicht zu*) bis 4 (*trifft völlig zu*) zu beantworten. Die interne Konsistenz war mit $\alpha = .86$ gut ($M = 3.14$; $SD = 0.56$; $ICC1 = .05$).

Umgang mit fehlenden Werten und methodisches Vorgehen

Aufgrund von Nichtbeantwortung einzelner Items kam es zu fehlenden Werten (bei keinem Item häufiger als 5 %, bei keinem Humoritem häufiger als 1 %), die mit Hilfe des Expectation-Maximization-Algorithmus imputiert wurden (vgl. Peugh & Enders, 2004).

Es wurde eine Analysestrategie gewählt, die der geschachtelten Datenstruktur gerecht wird (der Humor jeder Lehrkraft wurde von mehreren Lernenden beurteilt). Abgesehen von grundlegenden Item- und Skalenanalysen

wurden alle Analysen mit Hilfe von Mplus 6 (Muthén & Muthén, 2010) durchgeführt. Dabei behandelten wir alle Items als ordinalskaliert und wendeten entsprechend den Means-and-Variance-Adjusted-Weighted-Least-Squares-Schätzer (WLSMV-Schätzer) an.

Ergebnisse

Itemanalysen

Tabelle 1 zeigt Mittelwerte, Standardabweichungen, Schiefe und Intraklassenkorrelation ($ICC1$) aller Items der Konstruktionsversion der Skala zur Erfassung des Lehrerhumors. Fünf Items wiesen einige sehr ungünstige Kennwerte auf (Items 21, 22, 24, 25 und 26). Vor allem die sehr niedrigen Mittelwerte verweisen darauf, dass der darin angesprochene (aggressive) Humor sehr selten vorkommt und nicht wesentlich zur Differenzierung zwischen Lehrkräften beiträgt. Daher wurden diese Items verworfen. Bei allen Items waren die für die Erfassung von Lernumweltmerkmalen notwendigen moderaten bis großen Un-

Tabelle 1. Itemformulierungen und Itemkennwerte

| Nr. | Itemstamm: Unser [Fach-]Lehrer ... | <i>M</i> | <i>SD</i> | Schiefe | <i>ICC1</i> |
|-----|--|----------|-----------|---------|-------------|
| 1 | <i>verwendet Humor im Zusammenhang des Unterrichtsstoffs.</i> | 2.8 | 1.2 | 0.1 | .49 |
| 2 | <i>verwendet witzige Dinge zur Veranschaulichung oder als Beispiel.</i> | 2.4 | 1.1 | 0.4 | .40 |
| 3 | <i>erzählt uns Witze, die sich auf den Unterrichtsinhalt beziehen.</i> | 2.0 | 1.0 | 0.9 | .39 |
| 4 | <i>erzählt uns witzige Geschichten, die zum Unterricht passen.</i> | 2.1 | 1.0 | 0.7 | .36 |
| 5 | <i>spielt uns etwas vor oder führt etwas auf, um den Stoff zu veranschaulichen.</i> | 1.9 | 1.0 | 1.0 | .23 |
| 6 | <i>lässt uns Rollenspiele durchführen, um den Unterrichtsinhalt zu veranschaulichen.</i> | 2.4 | 1.1 | 0.3 | .44 |
| 7 | <i>verwendet eine lustige Sprache, um den Stoff zu erklären.</i> | 1.8 | 1.1 | 1.3 | .31 |
| 8 | <i>* bringt uns den Stoff auf humorvolle Art bei.</i> | 2.3 | 1.2 | 0.6 | .42 |
| 9 | <i>* verwendet lustige Beispiele im Unterricht.</i> | 2.2 | 1.0 | 0.6 | .40 |
| 10 | <i>erzählt uns Geschichten, die nichts mit dem Unterricht zu tun haben.</i> | 1.8 | 0.9 | 1.2 | .19 |
| 11 | <i>erzählt uns Witze, die mit dem Stoff nichts zu tun haben.</i> | 1.6 | 0.9 | 1.6 | .30 |
| 12 | <i>* schweift vom Thema ab und erzählt uns stattdessen witzige Geschichten.</i> | 1.4 | 0.7 | 2.1 | .22 |
| 13 | <i>macht sich im Unterricht über sich selbst lustig.</i> | 1.7 | 1.0 | 1.4 | .18 |
| 14 | <i>erzählt Geschichten, die für ihn selbst peinlich sind.</i> | 1.5 | 0.8 | 2.0 | .18 |
| 15 | <i>macht sich über sich selbst lustig, wenn er einen Fehler im Unterricht macht.</i> | 1.9 | 1.0 | 0.9 | .20 |
| 16 | <i>* lacht über sich selbst.</i> | 2.4 | 1.2 | 0.4 | .24 |
| 17 | <i>macht sich im Unterricht über Schüler lustig.</i> | 1.9 | 1.2 | 1.1 | .32 |
| 18 | <i>macht sich im Unterricht über die Intelligenz von Schülern lustig.</i> | 1.7 | 1.0 | 1.6 | .31 |
| 19 | <i>macht witzige Bemerkungen über das Privatleben oder persönliche Vorlieben von Schülern.</i> | 1.6 | 1.0 | 1.6 | .21 |
| 20 | <i>macht sich über den Kleidungsstil von Schülern lustig.</i> | 1.3 | 0.8 | 3.2 | .19 |
| 21 | <i>veralbert Schüler, weil sie ein Mädchen oder ein Junge sind.</i> | 1.2 | 0.6 | 3.8 | .21 |
| 22 | <i>macht witzige Bemerkungen über die Religion von Schülern.</i> | 1.1 | 0.4 | 6.0 | .21 |
| 23 | <i>benutzt Vorurteile, um witzig zu sein.</i> | 1.3 | 0.7 | 2.7 | .28 |
| 24 | <i>macht Witze über bestimmte Bevölkerungsgruppen (z. B. Ausländer).</i> | 1.2 | 0.7 | 3.7 | .35 |
| 25 | <i>macht Witze über bestimmte Religionen.</i> | 1.1 | 0.5 | 5.4 | .32 |
| 26 | <i>macht Witze über Homosexuelle (Schwule und Lesben).</i> | 1.1 | 0.5 | 5.0 | .20 |

Anmerkungen: $N = 890$ Schülerinnen und Schüler in 41 Schulklassen. $ICC1$ = Intraklassenkorrelation 1 (Anteil der Zwischenklassenvarianz an der Gesamtvarianz). Neu-konstruierte Items sind mit * gekennzeichnet, die übrigen Items sind Adaptionen aus dem Instrument von Frymier et al. (2008). Items der finalen Version des Fragebogens sind kursiv gedruckt.

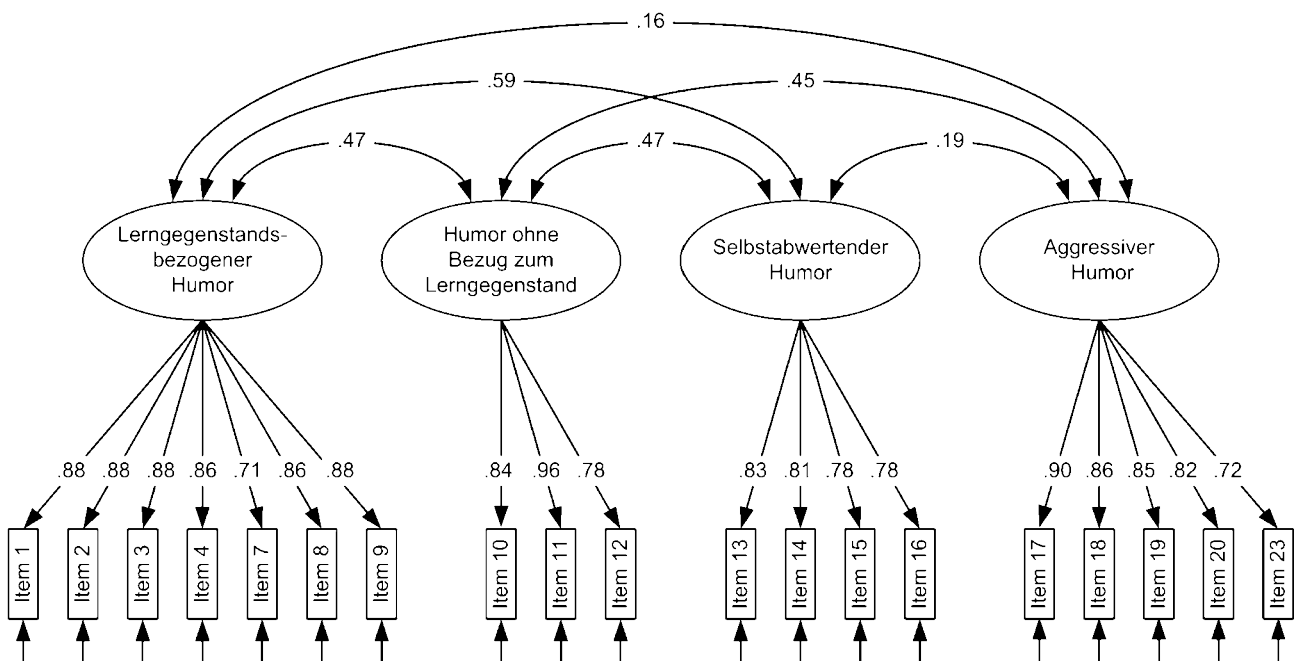


Abbildung 2. Ergebnisse von konfirmatorischen Faktorenanalysen des Lehrerhumors auf Schülerebene: Standardisierte Ladungen und Faktorkorrelationen des postulierten Modells mit vier Faktoren (alle $p < .05$).

terschiede zwischen Schulklassen bzw. Lehrkräften zu beobachten ($ICC1 = .18-.49$).

Faktorenstruktur des wahrgenommenen Lehrerhumors

Die Analyse der faktoriellen Validität (Hypothese 1) erfolgte wegen der auf Klassenebene eingeschränkten Stichprobe mit Hilfe eines gestuften Vorgehens (vgl. Hox, Maas & Brinkhuis, 2010): Im ersten Schritt wurde die postulierte vier-faktorielle Struktur des Lehrerhumors mit Hilfe von konfirmatorischen Faktorenanalysen auf Ebene der individuellen Schülerwahrnehmungen getestet, wobei die Standardfehler für die geschachtelte Datenstruktur korrigiert wurden (Mplus-Option „type is complex“). Um Hinweise zur Dimensionalität des wahrgenommenen Lehrerhumors auch auf Klassenebene zu erhalten, wurden im zweiten Schritt explorative Zwei-Ebenen-Faktorenanalysen mit Geomin-Rotation durchgeführt (vgl. Muthén & Asparouhov, 2010).

Das postulierte Vier-Faktoren-Modell, bei dem alle Items auf dem theoretisch zugehörigen Faktor laden, erbrachte in den konfirmatorischen Analysen auf Schülerebene eine akzeptable Passung ($\chi^2 = 451.8$; $df = 183$; $p < .001$; CFI = .95; RMSEA = .04). Die Items 5 und 6, deren Ladungen (.58 und .39) unterhalb der von Kline (2005) empfohlenen Schwelle von .60 lagen, wurden verworfen. Das reduzierte Modell wies ebenfalls eine akzeptable Datenpassung auf ($\chi^2 = 411.1$; $df = 146$; $p < .001$; CFI = .96; RMSEA = .05). Standardisierte Ladungen und

Faktorkorrelationen dieses Modells sind in Abbildung 2 dargestellt. Als Alternative zum postulierten Vier-Faktoren-Modell testeten wir ein Modell, bei dem alle Humortems auf einem einzigen Faktor laden. Dieses Modell war dem postulierten Modell deutlich unterlegen ($\chi^2 = 2386.9$; $df = 152$; $p < .001$; CFI = .64; RMSEA = .13).

In den anschließenden explorativen Zwei-Ebenen-Faktorenanalysen zur Dimensionalität des wahrgenommenen Lehrerhumors auf Klassenebene wurde eine Serie von Modellen mit variierender Faktorenzahl geschätzt, wobei die Anzahl der Faktoren auf Schülerebene – als Ergebnis des ersten Schritts – auf Vier fixiert wurde (siehe Tabelle 2).

Zusätzlich schätzten wir als Referenz für die Datenpassung ein freies Korrelationsmodell auf Klassenebene. Eine akzeptable Datenpassung war für die Modelle mit zwei oder mehr Faktoren auf Klassenebene zu beobachten, wobei sich diese mit steigender Faktorenzahl erkennbar verbesserte (zur Beurteilung der Datenpassung auf der zweiten Ebene ist vor allem der SRMR-Index auf Klassenebene relevant). Die Ladungsmuster der Modelle mit zwei oder drei Faktoren auf Klassenebene waren durch eine größere Zahl an Doppelladungen gekennzeichnet. Das Modell mit vier Faktoren auf Klassenebene wies dagegen nur eine Doppelladung auf. Insgesamt erwies sich das Vier-Faktoren-Modell damit als überlegen. Item 23, bei dem die genannte Doppelladung auftrat, sowie Item 7, das auf dem zugehörigen Klassenfaktor des *lerngegenstandsbezogenen Humors* eine zu geringe Ladung aufwies (.53), wurden eliminiert (vgl. Kline, 2005).

Tabelle 2. Ergebnisse von explorativen Zwei-Ebenen Faktorenanalysen des Lehrerhumors: Datenpassung für Modelle mit variierender Faktorenzahl auf Klassenebene

| Anzahl Faktoren | | SRMR | | | | | |
|-----------------|--------------|------|-----------------------|-----|-------|--------------|--------------|
| Schülerebene | Klassenebene | df | χ^2 ^a | CFI | RMSEA | Schülerebene | Klassenebene |
| 4 | frei | 101 | 226.5 | .97 | .04 | .03 | .00 |
| 4 | 1 | 253 | 435.5 | .96 | .03 | .03 | .22 |
| 4 | 2 | 235 | 368.4 | .97 | .03 | .03 | .06 |
| 4 | 3 | 218 | 355.3 | .97 | .03 | .03 | .04 |
| 4 | 4 | 202 | 344.1 | .97 | .03 | .03 | .02 |

Anmerkungen: $N = 890$ Schülerinnen und Schüler in 41 Schulklassen. Als Ergebnis der auf Schülerebene durchgeführten konfirmatorischen Faktorenanalysen wurde die Faktorenzahl dort auf 4 fixiert (vgl. Text). Die Analysen wurden mit Mplus 6 durchgeführt (Geomin-Rotation; WLSMV-Schätzer). ^a Alle $p < .001$

Tabelle 3. Ergebnisse von explorativen Zwei-Ebenen Faktorenanalysen des Lehrerhumors: Standardisierte Ladungen des Modells mit jeweils vier Faktoren auf Schüler- und Klassenebene

| Faktor | Item | Schülerebene | | | | Klassenebene | | | |
|--|------|--------------|------------|------------|------------|--------------|------------|-------------|-------------|
| | | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 1. Lerngegenstandsbezogener Humor | 1 | .79 | -.03 | -.04 | -.02 | .77 | .33 | .05 | -.04 |
| | 2 | .86 | -.04 | -.00 | -.06 | .78 | .29 | -.00 | -.01 |
| | 3 | .83 | -.03 | .06 | -.04 | .79 | .12 | .00 | .16 |
| | 4 | .77 | .05 | .03 | .07 | .89 | -.05 | -.15 | .20 |
| | 8 | .67 | .10 | -.13 | .06 | .76 | .33 | .01 | -.02 |
| | 9 | .71 | .12 | .03 | .02 | .70 | .26 | .06 | .12 |
| 2. Humor ohne Bezug zum Lerngegenstand | 10 | -.08 | .02 | -.01 | .91 | .07 | -.03 | .00 | .97 |
| | 11 | .03 | .02 | .01 | .83 | .17 | .04 | .10 | .80 |
| | 12 | .09 | -.06 | .10 | .62 | -.08 | .03 | .00 | 1.02 |
| 3. Selbstabwertender Humor | 13 | -.08 | .88 | .05 | -.02 | -.02 | .99 | .17 | .04 |
| | 14 | .03 | .70 | .14 | .03 | .02 | .62 | -.01 | .39 |
| | 15 | .04 | .78 | -.01 | -.06 | .04 | .97 | -.20 | -.04 |
| | 16 | .17 | .56 | -.10 | .12 | .10 | .87 | -.07 | .07 |
| 4. Aggressiver Humor | 17 | -.04 | .01 | .81 | .03 | -.01 | .12 | .98 | .02 |
| | 18 | .03 | -.10 | .89 | -.01 | -.09 | .02 | 1.00 | -.03 |
| | 19 | -.02 | .11 | .74 | -.01 | .15 | -.03 | .89 | .12 |
| | 20 | -.02 | .09 | .67 | .07 | .33 | -.22 | .77 | -.01 |

Anmerkungen: $N = 890$ Schülerinnen und Schüler in 41 Schulklassen. Die Analysen wurden mit Mplus 6 durchgeführt (Geomin-Rotation; WLSMV-Schätzer). Faktorladungen $> .40$ sind fett gedruckt (nach Jöreskog (1999), können Ladungen korrelierter Faktoren in Ausnahmefällen auch größer 1 sein). Die Itemformulierungen finden sich in Tabelle 1.

Tabelle 3 zeigt das Ladungsmuster für die favorisierte Lösung mit jeweils vier Faktoren auf Schüler- und Klassenebene mit den finalen 17 Items ($\chi^2 = 286.9$; $df = 148$; $p < .001$; CFI = .97; RMSEA = .03; SRMR Schülerebene = .03; SRMR Klassenebene = .02).¹ Dieses wies für alle Items sowohl auf Schüler- als auch auf Klassenebene hinreichend große Ladungen auf dem postulierten Faktor und hinreichend kleine Ladungen auf den übrigen Faktoren auf.

Skalenanalysen

Basierend auf den Ergebnissen der Faktorenanalysen bildeten wir vier Skalen mit insgesamt 17 Items, die die vier angenommenen Formen des Lehrerhumors reflektieren (entsprechend den Zuordnungen zu Faktoren in Tabelle 3). Die Subskala *lerngegenstandsbezogener Humor* erfasst mit sechs Items das Ausmaß, in dem die Lehrkraft den Lerngegenstand etwa durch humorige Beispiele verdeutlicht. Die Subskala *Humor ohne Bezug zum Lerngegenstand* erfasst mit drei Items, inwiefern die Lehrkraft Witziges im Unterricht zum Besten gibt, das jedoch nicht im Zusammenhang mit dem Unterrichtsstoff steht. Mit vier Items erfasst eine weitere Subskala *selbstabwertender Humor*, inwiefern die Lehrkraft über sich

¹ Auch eine konfirmatorische Faktorenanalyse auf Schülerebene mit den finalen 17 Items erbrachte eine gute Datenpassung ($\chi^2 = 399.3$; $df = 113$; $p < .001$; CFI = .96; RMSEA = .05)

selbst lachen kann und sich über sich selbst lustig macht. Die Subskala *aggressiver Humor* bildet schließlich mit vier Items das Ausmaß ab, in dem die Lehrkraft Humor auf Kosten der Schülerinnen und Schüler zeigt, etwa wenn sie sich über die Leistungen, Vorlieben oder Äußerlichkeiten von Schülerinnen und Schülern lustig macht.

Die internen Konsistenzen der vier Subskalen waren gut bis sehr gut (Tabelle 4). Auf Klassenebene zeigte sich durchwegs eine exzellente interne Konsistenz von Cronbachs $\alpha > .90$. Ähnliches gilt auch für das zweite berechnete Reliabilitätsmaß, die Intraklassenkorrelation *ICC2* – diese spezifiziert die Homogenität von Messungen, bei denen ein Merkmal mit Hilfe mehrerer individueller Fremdeinschätzungen (hier: zu Klassenmittelwerten aggregierte Schülerwahrnehmungen) erfasst wird (wobei Werte von *ICC2* $\geq .70$ eine gute Reliabilität indizieren; vgl. Lüdtke, Trautwein, Kunter & Baumert, 2006). Basierend auf den beobachteten großen Unterschieden zwischen Lehrkräften (*ICC1* $\geq .24$) zeigten sich hierbei durchgehend sehr gute Reliabilitätswerte der aggregierten Schülerwahrnehmungen (*ICC2* $\geq .87$). Die Trennschärfen der finalen Items in Bezug auf ihre zugehörigen Faktoren lagen auf Schülerebene im Bereich $r_{it} = .51\text{--}.82$ und auf Klassenebene im Bereich $r_{it} = .71\text{--}.97$.

Lerngegenstandsbezogener Humor, Humor ohne Bezug zum Lerngegenstand sowie selbstabwertender Humor korrelierten moderat positiv (vgl. Tabelle 4). Die Zusammenhänge mit *aggressivem Humor* waren in der Tendenz ebenfalls positiv, aber nicht in jedem Fall signifikant.

Konvergente und diskriminante Validität

Zur Überprüfung der konvergenten Validität (Hypothese 2) wurden die Korrelationen der vier Subskalen des Lehrerhumors mit den erfassten Außenmaßen simultan auf Schüler- und Klassenebene bestimmt, wobei die Variablen auf Schülerebene am jeweiligen Klassenmittelwert zentriert wurden (siehe Tabelle 5).

Erwartungsgemäß zeigten sich moderat positive Zusammenhänge des *lerngegenstandsbezogenen Lehrerhumors* mit dem wahrgenommenen Lehrerenthusiasmus und der wahrgenommenen Interessantheit des Unterrichts. Wie häufig der Fall, fiel das Muster der Zusammenhänge auf Klassenebene noch deutlicher aus (vgl. Marsh et al., 2012). Dies galt auch für die Zusammenhänge des *selbstabwertenden Lehrerhumors*. Für den *Humor ohne Bezug zum Lerngegenstand* war dagegen nur ein moderat positiver Zusammenhang mit der erlebten Interessantheit auf Klassenebene evident. *Aggressiver Lehrerhumor* korrelierte wie erwartet schwach bis moderat negativ mit den Schülerwahrnehmungen von Lehrerenthusiasmus und Interessantheit des Unterrichts.

Im Hinblick auf die diskriminante Validität (Hypothese 3) waren – wie erwartet – keine substanziellen Korrela-

Tabelle 4. Ergebnisse der Skalen- und Korrelationsanalysen auf Schüler- und Klassenebene

| Form des Lehrerhumors | M | SD | Schiefe | α | ICC1 | ICC2 | Bivariate Korrelationen | | |
|--|------|------|---------|----------|------|------|-------------------------|------|------|
| | | | | | | | 1 | 2 | 3 |
| Schülerebene | | | | | | | | | |
| 1. Lerngegenstandsbezogener Humor | 2.30 | 0.93 | 0.59 | .93 | .46 | .95 | | | |
| 2. Humor ohne Bezug zum Lerngegenstand | 1.58 | 0.71 | 1.50 | .81 | .28 | .89 | .12* | | |
| 3. Selbstabwertender Humor | 1.88 | 0.81 | 0.98 | .81 | .24 | .87 | .30* | .29* | |
| 4. Aggressiver Humor | 1.63 | 0.84 | 1.73 | .85 | .31 | .91 | -.04 | .32* | .20* |
| Klassenebene | | | | | | | | | |
| 1. Lerngegenstandsbezogener Humor | 2.30 | 0.67 | 0.43 | .98 | — | — | | | |
| 2. Humor ohne Bezug zum Lerngegenstand | 1.58 | 0.41 | 2.16 | .96 | — | — | .35 | | |
| 3. Selbstabwertender Humor | 1.88 | 0.43 | 0.64 | .93 | — | — | .53* | .51* | |
| 4. Aggressiver Humor | 1.63 | 0.49 | 1.41 | .93 | — | — | .21* | .44* | -.03 |

Anmerkungen: N = 890 Schülerinnen und Schüler in 41 Schulklassen. *ICC1* = Intraklassenkorrelation 1 (Anteil der Zwischenklassenvarianz an der Gesamtvarianz). *ICC2* = Intraklassenkorrelation 2 (Maß für die Reliabilität der Schulklassenmittelwerte der einzelnen Schülerwahrnehmungen). Die Korrelationen wurden simultan auf beiden Ebenen mit Hilfe von Mplus 6 bestimmt (WLSMV-Schätzer; Variablen auf Schülerebene wurden am jeweiligen Klassenmittelwert zentriert). * $p < .05$.

Tabelle 5. Korrelationen zwischen Lehrerhumor und Außenkriterien auf Schüler- und Klassenebene

| Form des Lehrerhumors | ENT | INT | ZEI | HOS | SWS |
|-------------------------------------|-------|-------|-------|-------|------|
| Schülerebene | | | | | |
| Lerngegenstandsbezogener Humor | .21* | .45* | .14* | .04 | .09* |
| Humor ohne Bezug zum Lerngegenstand | -.02 | .00 | -.23* | .06 | .00 |
| Selbstabwertender Humor | .10* | .28* | -.03 | .03 | .03 |
| Aggressiver Humor | -.13* | -.10* | -.23* | .06 | .01 |
| Klassenebene | | | | | |
| Lerngegenstandsbezogener Humor | .40* | .49* | .11 | -.55* | -.04 |
| Humor ohne Bezug zum Lerngegenstand | .14 | .42* | -.22 | -.16 | -.05 |
| Selbstabwertender Humor | .52* | .78* | .21 | -.03 | .29 |
| Aggressiver Humor | -.35* | -.03 | .04 | -.08 | -.32 |

Anmerkungen: $N = 890$ Schülerinnen und Schüler in 41 Schulklassen. Die Korrelationen wurden simultan auf beiden Ebenen mit Hilfe von Mplus 6 bestimmt (WLSMV-Schätzer; Variablen auf Schülerebene wurden am jeweiligen Klassenmittelwert zentriert). ENT = Enthusiasmus der Lehrkraft. INT = Interessantheit des Unterrichts. ZEI = Zeitznutzung im Unterricht. HOS = Humororientierung der Schülerinnen und Schüler. SWS = Selbstwirksamkeit der Schülerinnen und Schüler. * $p < .05$.

tionen mit der wahrgenommenen Zeitznutzung im Unterricht, der Humororientierung der Schülerinnen und Schüler oder der Selbstwirksamkeitsüberzeugung der Schülerinnen und Schüler evident (vgl. Tabelle 5). Eine Ausnahme davon stellte lediglich ein moderat negativer Zusammenhang auf Klassenebene zwischen der Humororientierung der Schülerinnen und Schüler und *lerngegenstandsbezogenem Humor* dar.

Diskussion

Das Ziel der vorliegenden Arbeit war es, auf der Basis des theoretischen Modells von Martin et al. (2003) und des für den US-amerikanischen Hochschulforschungsbereich konstruierten Fragebogens von Frymier et al. (2008) ein deutschsprachiges Instrument zur Erfassung des Humors von Lehrkräften zu entwickeln und zu validieren. Zentrale Stärken des vorgelegten Instruments und der berichteten Studie sind die Einnahme einer multidimensionalen Perspektive auf Lehrerhumor, die Nutzung von Schülerwahrnehmungen zu dessen Operationalisierung und die systematische Berücksichtigung von Schüler- als auch Lehrkräfteebene.

In Bezug auf die faktorielle Validität des entwickelten Instruments (Hypothese 1) lieferten die Analysen sehr klare Evidenz für das postulierte Modell mit vier Formen des Lehrerhumors. Konfirmatorische Faktorenanalysen erbrachten eine gute Datenpassung dieses Modells auf Ebene der individuellen Schülerwahrnehmungen. Aus seiner Überlegenheit gegenüber einem Ein-Faktor-Modell lässt sich folgern, dass Schülerinnen und Schüler in der Lage sind, verschiedene Formen des Lehrerhumors differenziert wahrzunehmen. Die Ergebnisse von explorativen Zwei-Ebenen-Faktorenanalysen standen auch auf Klassenebene im Einklang mit den vier angenommenen Humorformen. Diese Ergebnisse bestätigen die bisherigen Befunde, die für die Annahme einer differenzierten Humorwahr-

nehmung bei Studierenden sprachen (Frymier et al., 2008; Wanzer et al., 2010) – zugleich erweitern sie den Altersbereich für die Gültigkeit dieser Annahme nach unten.

Daneben zeigten sich aber auch einige Unterschiede zu den Befunden aus US-amerikanischen Colleges (z. B. Frymier et al., 2008; Wanzer & Frymier, 1999; Wanzer et al., 2006; Wanzer et al., 2010). So fand sich in der vorliegenden Studie – wie erwartet – keine Evidenz dafür, dass schulische Lehrkräfte in nennenswertem Umfang besonders *aggressiven Humor* einsetzen, der auf Stereotypen basiert und bestimmte Personengruppen herabsetzt (vgl. Hill, 1988; Neuliep, 1991). Weiterhin fielen die Zusammenhänge zwischen den vier Humorformen teilweise sogar etwas niedriger aus als in den Studien von Frymier et al. (2008) und Wanzer et al. (2010), auch wenn das gefundene Korrelationsmuster insgesamt sehr gut mit früheren Befunden korrespondiert. Dieses unterstreicht nochmals, dass Schülerinnen und Schüler ebenso wie Studierende sehr gut zwischen verschiedenen Formen des Humors von Lehrpersonen differenzieren können.

Die auf der Basis des validierten Vier-Faktoren-Modells gebildeten Skalen verfügen über eine gute bis sehr gute Reliabilität (Cronbachs α und ICC2), die den von Frymier et al. (2008) identifizierten, guten Kennwerten mindestens entsprechen. Insbesondere auf Klassenebene war die Reliabilität exzellent (die Schülerinnen und Schüler einer individuellen Lehrkraft beurteilten deren Humor mit angemessen kleinem Messfehler). Die Messgenauigkeiten lagen damit in einem Bereich, der eine Verwendung des Instruments im Rahmen einer Individualdiagnostik von Lehrkraftmerkmalen als denkbar erscheinen lässt – zukünftige Untersuchungen mit den dafür noch nötigen empirischen Schritten (z. B. Normierung) erachten wir deshalb als lohnend.

Evidenz für die konvergente Validität des Verfahrens (Hypothese 2) lieferten Korrelationen der verschiedenen

Formen des Lehrerhumors mit Schülerwahrnehmungen des Lehrerenthusiasmus und der Interessantheit des Unterrichts. Dabei zeigten sich differentielle Zusammenhangsmuster: Unterschiedliche Humorformen standen in unterschiedlichem Zusammenhang mit den Außenmaßen. Dies darf als weiterer Beleg für die Angemessenheit der Unterscheidung verschiedener Humorformen gewertet werden. Hypothesenkonform waren positive Zusammenhänge der von Lernenden als angemessen erlebten Humorformen sowie negative Zusammenhänge des als unangemessen erlebten *aggressiven Humors* zu beobachten (vgl. Wanzer et al., 2006). Auch die positiven Zusammenhänge zwischen *selbstabwertendem Lehrerhumor* und dem Enthusiasmus der Lehrkraft sowie der Interessantheit des Unterrichts stehen im Einklang mit bisherigen Ergebnissen, obwohl diese Art des Humors auf Kosten des Humorproduzenten geht und negative Auswirkungen für die betreffenden Lehrkräfte haben könnte (vgl. Wanzer et al., 2010). Von Lernenden wird jedoch auch *selbstabwertender Lehrerhumor* als unterhaltend wahrgenommen und als angemessen bewertet (Booth-Butterfield & Wanzer, 2010). Die bei diesen Analysen auf Klassenebene teilweise zu beobachtenden großen Zusammenhänge sind dabei nicht untypisch für die Aggregatebene der geteilten Schülerwahrnehmungen (auf der Messfehler deutlich reduziert sind), liegen in einem Bereich, der aus bisherigen Studien mit unidimensional erfasstem Lehrerhumor bereits bekannt sind, und stellen die Abgrenzbarkeit der beteiligten Merkmale nicht in Frage (vgl. Dresel et al., 2014).

Für die diskriminante Validität des Instruments sprachen die wie erwartet nicht-signifikanten oder nur kleinen Zusammenhänge der verschiedenen Humorformen mit der Zeitnutzung im Unterricht sowie der Humororientierung und der Selbstwirksamkeit der Schülerinnen und Schüler (Hypothese 3). Die schwach negativen Zusammenhänge des *Humors ohne Bezug zum Lerngegenstand* und des *aggressiven Humors* mit der Zeitnutzung, die sich auf der Schülerebene fanden, könnten daraufhin deuten, dass diese Humorformen von den Lernenden eher als Zeitverschwendung empfunden werden und damit wichtige Lehrfunktionen des Unterrichts nicht unterstützen (vgl. Klauer, 1985). Auch sonst lagen die wenigen signifikanten Zusammenhänge überwiegend auf der Ebene der (von der geteilten Wahrnehmung in der Klasse abweichenden) Individualwahrnehmung der Schülerinnen und Schüler und beeinträchtigten die Validität des Verfahrens auf Ebene der Klasse damit nicht. Die einzige Ausnahme davon war eine moderat negative Korrelation auf Klassenebene zwischen der Humororientierung der Schülerinnen und Schüler und dem wahrgenommenen *lerngegenstandsbezogenen Lehrerhumor*. Da die Humororientierung der Schülerinnen und Schüler auf Klassenebene jedoch nur eine sehr geringe Varianz aufwies ($ICC1 = .03$), dürfte dieser Zusammenhang nur von geringer substantieller Bedeutung sein.

Trotz des insgesamt sehr eindeutigen Ergebnismusters, das für die psychometrische Qualität des Instruments spricht, weist die vorgestellte Studie auch einige Limitationen auf. Erstens ist die Generalisierbarkeit der Befunde eingeschränkt, da in der vorliegenden Untersuchung nur Gymnasiasten in Bezug auf das Fach Deutsch untersucht wurden. Lehrkräfte können Humor jedoch prinzipiell in allen Fächern und Schularten in seinen verschiedenen Facetten einsetzen und dürften dies auch tun (wenngleich in variierender Intensität). Zudem sind Kinder in einem Alter von 11 bis 12 Jahren üblicherweise bereits in der Lage, komplexen Humor mit konzeptueller Inkongruenz und linguistischer Ambiguität zu verstehen (Bergen, 1998). Deshalb darf man für die Sekundarstufe insgesamt optimistisch sein, dass der Fragebogen dort erfolgreich eingesetzt werden kann. Limitierend ist, zweitens, die auf Klassenebene vergleichsweise kleine Stichprobe, die den Einsatz von explorativen anstelle von (noch angemesseneren) konfirmatorischen Zwei-Ebenen-Faktorenanalysen bedingte. Deren Ergebnisse waren jedoch vergleichsweise eindeutig, so dass die Zweifel an einer Struktur mit vier Formen des Lehrerhumors auf Klassenebene nicht allzu groß ausfallen dürften. Eine dritte Limitation liegt schließlich darin, dass zur Überprüfung der konvergenten Validität des Schülerfragebogens zum Lehrerhumor ebenfalls nur die Schülerperspektive herangezogen wurde, die Perspektive der Lehrkräfte selbst aber unberücksichtigt blieb. Auch um Diskrepanzen zwischen Selbst- und Fremdwahrnehmung des Lehrerhumors zu analysieren, wäre es wünschenswert, zukünftig mehrere Perspektiven zu erfassen (vgl. Clausen, 2002).

Auch wenn die genannten Limitationen auf noch offene Fragestellungen verweisen, belegen die vorgelegten Ergebnisse, dass das für den US-amerikanischen Hochschulbereich konstruierte Originalinstrument von Frymier et al. (2008) erfolgreich auf den deutschsprachigen Schulkontext adaptiert wurde. In der Zusammenschau der Ergebnisse stellt sich der „Fragebogen zur Erfassung des Humors von Lehrkräften aus Schülersicht“ (HUMLAS) als reliables und valides Instrument dar, das bisherige Schwachstellen bei der Erfassung des Lehrerhumors überwindet – insbesondere seine eindimensionale Konzeption und seine Operationalisierung durch Lehrerselbstberichte. Damit kann ein bislang noch zu selten untersuchter Aspekt des Unterrichts und des professionellen Handelns von Lehrkräften künftig differenziert und ökonomisch erfasst werden.

Literatur

- Banas, J. A., Dunbar, N., Rodriguez, D. & Liu, S.-J. (2011). A review of humor in educational settings: Four decades of research. *Communication Education*, 60, 115–144.
- Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U. et al. (2008). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und*

- die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente (Materialien aus der Bildungsforschung, Nr. 83). Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J., Gruehn, S., Heyn, S., Köller, O. & Schnabel, K.-U. (1997). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). Dokumentation, Band 1. Skalen Längsschnitt I, Welle 1–4*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Bergen, D. (1998). Development of the sense of humor. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 329–358). New York, NY: Mouton de Gruyter.
- Berlyne, D. E. (1960). *Conflict, arousal and curiosity*. New York, NY: Mc Graw-Hill.
- Berlyne, D. E. (1963). Complexity and incongruity variables as determinants of exploratory choice and evaluative ratings. *Canadian Journal of Psychology*, 17, 274–290.
- Booth-Butterfield, S. & Booth-Butterfield, M. (1991). Individual differences in the communication of humorous messages. *Southern Communication Journal*, 56, 205–217.
- Booth-Butterfield, S. & Wanzer, M. B. (2010). Humor and communication in instructional contexts: Goal-oriented communication. In D. L. Fassett & J. T. Warren (Eds.), *The SAGE handbook of communication and instruction* (pp. 221–239). Thousand Oaks, CA: SAGE.
- Clausen, M. (2002). *Qualität von Unterricht – Eine Frage der Perspektive?* Waxmann: Münster.
- Davies, C. (1998). The dog that didn't bark in the night: A new sociological approach to the cross-cultural study of humor. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 293–306). New York, NY: Mouton de Gruyter.
- Dresel, M., Bieg, S., Fasching, M., Steuer, G., Nitsche, S. & Dickhäuser, O. (2014). Humor von Lehrkräften in der Schülerwahrnehmung: Abgrenzung von Lehrerenthusiasmus und Zusammenhänge mit Dimensionen des Unterrichts. *Psychologie in Erziehung und Unterricht*, 61, 56–74.
- Ditton, H. & Merz, D. (1999). *Interessantheit des Unterrichts*. Zugriff am 03.05.2012. Verfügbar unter http://www.quassu.net/Sch_fgb1.pdf
- Frymier, A. B., Wanzer, M. B. & Wojtaszczyk, A. M. (2008). Assessing students' perceptions of inappropriate and appropriate teacher humor. *Communication Education*, 57, 266–288.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität*. Seelze-Velber: Kallmeyer/Klett.
- Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41, 111–127.
- Hill, D. J. (1988). *Humor in the Classroom. A handbook for teachers (and other entertainers!)* Springfield, IL: Thomas.
- Hox, J. J., Maas, C. J. M. & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica*, 64, 157–170.
- Jerusalem, M. & Satow, L. (1999). Schulbezogene Selbstwirksamkeit. In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen* (S. 18–19). Berlin: Institut für Psychologie, Freie Universität Berlin.
- Jöreskog, K. G. (1999). *How large can a standardized coefficient be?* Zugriff am 30.07.2014. Verfügbar unter <http://www.ssicentral.com/lisrel/techdocs/HowLargeCanaStandardizedCoefficientbe.pdf>
- Klauser, K. J. (1985). Framework for a theory of teaching. *Teaching and Teacher Education*, 1, 5–17.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.
- Kubie, L. S. (1971). The destructive potential of humor in psychotherapy. *American Journal of Psychiatry*, 127, 37–42.
- Kuiper, N. A., Martin, R. A. & Olinger, L. J. (1993). Coping humour, stress, and cognitive appraisals. *Canadian Journal of Behavioural Science*, 25, 81–96.
- Kunter, M., Tsai, Y.-M., Klusmann, U., Brunner, M., Krauss, S. & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, 18, 468–482.
- LaFave, L., Haddad, J. & Maesen, W. A. (1996). Superiority, enhanced self-esteem, and perceived incongruity humor theory. In A. J. Chapman & H. C. Foot (Eds.), *Humor and laughter: Theory research and applications* (pp. 63–91). New Brunswick, NJ: Transaction.
- Lefcourt, H. M. (2001). *Humor: The psychology of living buoyantly*. New York, NY: Kluwer Academic Publishers.
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2006). Analyse von Lernumwelten: Ansätze zur Bestimmung der Reliabilität und Übereinstimmung von Schülerwahrnehmungen. *Zeitschrift für Pädagogische Psychologie*, 20, 85–96.
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S. et al. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, 47, 106–124.
- Martin, R. A. (2003). Sense of humor. In S. J. Lopez & C. R. Snyder (Eds.), *Positive psychological assessment: A handbook of models and measures* (pp. 313–326). Washington, DC: American Psychological Association.
- Martin, R. A. (2007). *The psychology of humor: An integrative approach*. Oxford: Elsevier Academic Press.
- Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J. & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality*, 37, 48–75.
- Muthén, B. & Asparouhov, T. (2010). Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 15–40). New York, NY: Routledge.
- Muthén, L. K. & Muthén, B. O. (2010). *Mplus (Version 6) [Computer Software]*. Los Angeles, CA: Muthén & Muthén.
- Neuliep, J. W. (1991). An examination of the content of high school teachers' humor in the classroom and the development of an inductively derived taxonomy of classroom humor. *Communication Education*, 40, 343–355.
- Nevo, O., Nevo, B. & Yin, J. L. S. (2001). Singaporean humor: A cross-cultural, cross-gender comparison. *Journal of General Psychology*, 128, 143–156.
- Peugh, J. L. & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, 74, 525–556.
- Riðland, B. (2002). *Humor und seine Bedeutung für den Lehrerberuf*. Bad Heilbrunn: Klinkhardt.
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. *Zeitschrift für Pädagogik*, 44, 28–53.

- Turner, J. C., Midgley, C., Meyer, D. K., Gheen, M., Anderman, E. M., Kang, Y. et al. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: A multimethod study. *Journal of Educational Psychology*, 94, 88–106.
- Wanzer, M. B. & Frymier, A. B. (1999). The relationship between student perceptions of instructor humor and students' reports of learning. *Communication Education*, 48, 48–62.
- Wanzer, M. B., Frymier, A. B. & Irwin, J. (2010). An explanation of the relationship between instruction humor and student learning: Instructional humor processing theory. *Communication Education*, 59, 1–18.
- Wanzer, M. B., Frymier, A. B., Wojtaszczyk, A. M. & Smith, T. (2006). Appropriate and inappropriate uses of humor by teachers. *Communication Education*, 55, 178–196.
- Zillman, D. (1983). Disparagement humor. In P. E. McGhee & J. H. Goldstein (Eds.), *Handbook of humor research* (Vol.1, pp. 85–108). New York, NY: Springer.
- Ziv, A. (1979). The teacher's sense of humor and the atmosphere in the classroom. *School Psychology International*, 1, 21–23.

Dr. Sonja Bieg
Prof. Dr. Markus Dresel

Universität Augsburg
Lehrstuhl für Psychologie
Universitätsstraße 10
86135 Augsburg
E-mail: sonja.bieg@phil.uni-augsburg.de