

**Empirische Sonderpädagogik**, 2015, Nr. 4, S. 283-299  
ISSN 1869-4845 (Print) · ISSN 1869-4934 (Internet)

## Nur ein Test? – Auswirkungen nicht unterstützenden Testleiterverhaltens bei Intelligenztestungen auf den Selbstwert

Nantje Otterpohl<sup>1</sup> & Malte Schwinger<sup>2</sup>

<sup>1</sup> Universität Bielefeld

<sup>2</sup> Universität Witten/Herdecke

### Zusammenfassung

Intelligenztestungen gehören zum beruflichen Alltag in der psychologischen und sonderpädagogischen Praxis. Während verschiedene Studien darauf hinweisen, dass bei der Durchführung von Intelligenztests die Art der Testinstruktion sowie das Testleiterverhalten einen bedeutsamen Einfluss auf die Testangst und die Testleistung haben, wurde bisher kaum untersucht, inwiefern diese Faktoren auch einen Einfluss auf das Selbstwertgefühl nach der Testung haben. Zur Untersuchung dieser Frage wurden mit  $N = 31$  Kindern der fünften Jahrgangsstufe standardisierte Einzeltestungen durchgeführt. Hierbei wurden die Kinder auf zwei Experimentalgruppen (unterstützende vs. nicht unterstützende Intelligenztestung) und eine Kontrollgruppe aufgeteilt. Bei den Kindern der nicht unterstützenden Experimentalgruppe zeigte sich im Vergleich zu den anderen beiden Gruppen eine signifikante Reduktion des situationalen Selbstwerts direkt im Anschluss an die Testung sowie des allgemeinen Selbstwerts zwei Wochen später. Die Effekte der Testbedingung erwiesen sich dabei als unabhängig von der erzielten Testleistung. Die Ergebnisse verdeutlichen, wie wichtig unterstützendes Testleiterverhalten bei der Durchführung von Intelligenztestungen ist, um nicht nur eine valide Leistungseinschätzung zu ermöglichen, sondern auch die emotionalen Kosten für die getesteten Kinder möglichst gering zu halten.

Schlüsselwörter: Intelligenztestung, Testleiterverhalten, Testinstruktion, Selbstwert, Emotionen

### Just a test? Effects of non-supportive intelligence instructions on children's self-esteem

#### Abstract

Intelligence tests are part of the daily work of psychotherapists and in special needs education. While a number of studies have demonstrated that in conducting intelligence tests the specific test instructions have a substantial impact on test anxiety and test performance, it has been seldom investigated in how far these factors also affect children's self-esteem after testing. To address this question, we tested  $N = 31$  fifth-grade children in a standardized, individual setting. Children were assigned to two experimental groups (supportive vs. non-supportive intelligence test instruction) and a control group. Compared to the other two groups, children of the non-

supportive condition showed significant reductions in situational self-esteem directly after test completion and also in overall self-esteem two weeks later. The effects were independent from children's actual test performance. The findings emphasize that supportive test instructions are not only important for valid assessments of test performance, but also to reduce the emotional costs for the children tested.

Keywords: intelligence test, test administration, test instruction, self-esteem, emotions

Individuelle Testungen der allgemeinen Intelligenz stellen in der psychologischen und sonderpädagogischen Praxis ein unverzichtbares Hilfsmittel bei der Diagnostik und Entscheidungsfindung dar. Die Gründe, einen Intelligenztest durchzuführen, können vielfältig sein und umfassen beispielsweise Fragen der Schullaufbahnenempfehlung, die Abklärung einer möglichen Hochbegabung, die Diagnostik von Lern- und Leistungsstörungen oder die Feststellung eines sonderpädagogischen Förderbedarfs. Angesichts der hohen Bedeutsamkeit von Intelligenztestungen in der Einzelfalldiagnostik wurden bereits früh verstärkte Anstrengungen unternommen, die Ergebnisse verschiedener Teilnehmerinnen und Teilnehmer eines Intelligenztests vergleichbar zu halten. Als eines der wichtigsten Mittel der Sicherstellung der Vergleichbarkeit wurden standardisierte Testinstruktionen eingeführt, die heutzutage in jedem Manual gängiger Testverfahren enthalten sind.

Trotz des hohen Grades an Standardisierung gibt es jedoch immer noch gewissen Spielraum für das Verhalten von Testleiterinnen und Testleitern während der Testung. Zudem kann nicht mit Sicherheit davon ausgegangen werden, dass sich Testleiterinnen und Testleitern in der diagnostischen Praxis durchgehend gemäß den standardisierten Instruktionvorgaben verhalten. Befunde verschiedener Studien weisen darauf hin, dass die Art der Testinstruktion sowie das Testleiterverhalten während der Testung einen bedeutsamen Einfluss auf die Angst, im Test zu versagen, sowie die Leistung im jeweiligen Test haben (z.B. Meijer & Oostdam, 2011; Sarason, 1961; Sorrentino, Brouwers, Hanna & Roney, 1996). Demgegenüber wurde bislang kaum unter-

sucht, inwiefern bestimmtes instruktionales Testleiterverhalten – unabhängig von der Leistung im Test – negative Konsequenzen für das situationale Selbstwertgefühl von Probanden hat, welche möglicherweise über die Testung hinaus auf den allgemeinen Selbstwert und das Wohlbefinden im Alltag generalisiert werden. Dieses Forschungsdefizit betrifft mit Blick auf die hohe Zahl an Kindern, die vor der Durchführung eines Intelligenztests bereits eine lange Geschichte schulischer Misserfolgsereignisse aufweisen, und angesichts der Tatsache, dass Schule neben der Vermittlung fachlicher Kompetenzen auch die Stärkung des individuellen Selbstwertgefühls und der Persönlichkeitsentwicklung zum Ziel hat (Bless & Mohr, 2007), eine Frage, welche von zentraler Bedeutung für die tägliche praktische Arbeit in sonderpädagogischen und psychotherapeutischen Settings ist. In der vorliegenden Arbeit werden daher die Befunde einer experimentellen Studie berichtet, in der die Effekte unterschiedlichen Testleiterverhaltens auf das situationale Selbstwertgefühl unmittelbar nach einer Intelligenztestung sowie den allgemeinen Selbstwert zwei Wochen nach der Testung untersucht wurden.

### *Testleiterverhalten bei Intelligenztestungen*

Die Mehrzahl der Studien, die sich mit Auswirkungen des Testleiterverhaltens bei Intelligenztestungen befassen, haben die Effekte verschiedener Testinstruktionen (beruhigend vs. neutral vs. verunsichernd; z.B. Petty & Harrell, 1977) und unterschiedlicher Arten von Feedback (positiv vs. negativ; z.B. Sarason, 1961) in den Blick genom-

men. Als Outcome-Variable wurde in der Regel auf die Testleistung fokussiert, wobei einige Studien zusätzlich die Testangst der Probanden als Mediatorvariable (z.B. Sassenrath & Kight, 1965) und verschiedene Persönlichkeitsmerkmale wie z.B. dispositionelle Testangst als Moderatorvariable (Meijer & Oostdam, 2007) untersucht haben. Wenngleich die Befunde im Detail leicht variieren, so zeigt sich doch im Gesamtbild ein deutlicher Beleg für die erwarteten Effekte verschiedener Instruktionen und Feedbackarten bei Intelligenztestungen. Sassenrath und Kight (1965) etwa zeigten, dass durch negative bzw. positive Rückmeldungen während einer Lernaufgabe die Prüfungsangst und die Leistung der Probanden beeinflusst werden kann. Motivierende und beruhigende Rückmeldungen führten in der Studie bei gleicher Fähigkeit der Probanden zu besseren Leistungen als negatives und verunsicherndes Feedback. Petty und Harrell (1977) konnten darüber hinaus nachweisen, dass durch festgelegte Instruktionen die Ergebnisse von IQ-Tests deutlich gesteigert werden können, wenn diese Instruktionen Programme zur Motivationssteigerung, zur Testerfahrung oder zur Reduzierung von Testangst beinhalten (für ähnliche Befunde siehe z.B. Meijer & Oostdam, 2007; Sorrentino et al., 1996). In Bezug auf mögliche Moderatoreffekte verschiedener Persönlichkeitsmerkmale stellten verschiedene Autoren fest, dass hoch ängstliche Versuchspersonen in Fähigkeitstests deutlich schlechter abschneiden als weniger ängstliche Versuchspersonen, wenn sie vorab eine Instruktion erhalten, die eine persönliche Bedrohung für ihren Selbstwert impliziert (z.B. Heisler & Schill, 1972; Meijer & Oostdam, 2011; Sarason, 1961). Bei Betrachtung der oben referierten Studien fällt auf, dass die Auswirkungen verschiedener Testinstruktionen fast ausschließlich in Bezug auf die Testleistung sowie leistungsnahe affektive Outcomes untersucht wurden, mögliche Konsequenzen für allgemeinere Aspekte der psychischen Anpassung wie den allgemeinen Selbstwert

und das subjektive Wohlbefinden hingegen wenig bis gar keine Beachtung fanden.

### ***Auswirkungen des Testleiterverhaltens auf das Selbstwertgefühl***

Ein großer Teil der in der Praxis administrierten Intelligenztests wird mit lernschwachen Kindern durchgeführt, die aufgrund einer Teilleistungsschwäche (Lese-Rechtsschreibstörung, Dyskalkulie) oder anderer spezifischer Ursachen eine Reihe von schulischen Misserfolgen erlebt haben. Diese werden häufig in einer für den eigenen Selbstwert ungünstigen Weise attribuiert, vor allem wenn Kinder die eigene mangelnde Begabung als grundsätzlichen Verursachungsfaktor ihrer Misserfolge benennen (Covington, 1992; Fischbach, Mähler, Schuchardt & Hasselhorn, 2010). Als Folge davon nehmen viele dieser Schülerinnen und Schüler Klassenarbeiten und jegliche Arten von Tests als potentiell selbstwertbedrohliche Bewertungssituationen wahr. Jeder Misserfolg hat dann negative Auswirkungen auf das Selbstwertgefühl und die motivationale Einstellung zum Lernen im Allgemeinen. Wie verschiedene Studien zeigen, kann solch ein Mechanismus schnell in eine Art Abwärtsspirale münden, bei der sich negative Erwartungen und tatsächliche Misserfolge reziprok verstärken (z.B. Marsh & Craven, 2006).

Vor diesem Hintergrund ist anzunehmen, dass auch die psychosozialen Erfahrungen, die ein (lernschwaches) Kind während einer Intelligenztestung macht, einen bedeutsamen Einfluss auf das Selbstwertgefühl in der spezifischen Testsituation (und darüber hinaus) haben können. In der vorliegenden Arbeit wird unterschieden zwischen dem allgemeinen Selbstwert als einer affektiven, globalen und zeitlich stabilen Einstellung einer Person gegenüber sich selbst (Rosenberg, 1965) und dem situationalen Selbstwert, der laut Heatherston und Polivy (1991) die auf den Moment bezogene und stärker fluktuierende Variante des ei-

genen Selbstwerts ausmacht. Effekte verschiedener Testinstruktionen und Feedbackarten auf den situationalen Selbstwert wurden bereits in verschiedenen Studien untersucht (z.B. Greenberg & Pyszczynski, 1985; Guay, Delisle, Fernet, Julien & Senécal, 2008). Allerdings wurden in diesen Untersuchungen vorwiegend Erwachsenenstichproben adressiert und lediglich die Art des Feedbacks, nicht jedoch das konkrete Testleiterverhalten variiert. Zudem wurde der situationale Selbstwert nur einmalig am Ende der Testung erfasst, weshalb keine Prüfung einer Veränderungshypothese möglich war. Des Weiteren wurde in keiner der genannten Studien ein möglicher Transfereffekt auf den allgemeinen Selbstwert untersucht. Theoretisch können Effekte unterschiedlichen Testleiterverhaltens zum einen indirekt über die tatsächliche Testleistung vermittelt werden, was der Argumentation früherer Studien zu Testangst und Leistung entsprechen würde (Meijer & Oostdam, 2011). Zum anderen kann das Testleiterverhalten aber auch direkte Effekte auf den situationalen Selbstwert ausüben, die unabhängig von der eigentlichen Testleistung wären. Von besonderer Bedeutung dürften diesbezüglich das Ausmaß an autonomieunterstützender Instruktion und Feedback sowie die Herstellung einer eher lern- statt leistungszielorientierten Testsituation sein.

*Erleben von Autonomie und Kompetenz.* Gemäß der Selbstbestimmungstheorie von Deci und Ryan (1985) stellen das Erleben von Autonomie und Kompetenz zentrale psychologische Grundbedürfnisse eines Menschen dar. Eine nachhaltige Befriedigung dieser Bedürfnisse trägt demnach zu einem gesunden Selbstwertgefühl und hohem psychischen Wohlbefinden bei (Ryan & Deci, 2000). In diesem Zusammenhang konnte gezeigt werden, dass autonomieunterstützendes Instruktionsverhalten von Lehrkräften und Testleiterinnen und Testleitern sich positiv auf die intrinsische Motivation, die Selbstwirksamkeit und den Selbstwert von Schülerinnen und Schülern jeden Alters auswirkt (Reeve & Jang, 2006). Be-

sondere Kennzeichen eines solchen autonomieunterstützenden Instruktionsstils beziehen sich z.B. auf Responsivität gegenüber Fragen und Lösungsvorschlägen von Schülerinnen und Schülern, Gewährung ausreichender Redezeit, Anerkennung alternativer Lösungswege oder die Bereitstellung von Auswahlalternativen. Hingegen fühlen sich Schülerinnen und Schüler in ihrer Autonomie zumeist nicht unterstützt, wenn der Lehrende kaum Rückfragen zulässt, keinerlei Feedback gibt, Wahlmöglichkeiten und Mitbestimmung im Unterricht einschränkt oder rigide Vorgaben in Bezug auf die zu erfüllende Stoffmenge macht (Reeve & Jang, 2006).

Das Grundbedürfnis nach Kompetenzerleben hängt ebenfalls stark mit dem Instruktionsverhalten zusammen. Zu hohe Erwartungen von Lehrkräften, inkonsistentes bzw. ausbleibendes Feedback oder stabile Fähigkeitszuschreibungen können das Erleben von Kompetenz auf Seiten der Schülerinnen und Schüler deutlich beeinträchtigen (Hattie, 2009; Jussim & Harber, 2005). Der Großteil der Studien zum Autonomie- und Kompetenzerleben von Kindern wurde im schulischen Setting durchgeführt (z.B. Reeve & Jang, 2006). Die hierbei gewonnenen Erkenntnisse können jedoch unserer Ansicht nach zu gewissen Teilen auf den Kontext des Testleiterverhaltens während einer Intelligenztestung übertragen werden. Sicherlich ist die Unterstützung von Autonomie- und Kompetenzerleben in einer Testsituation nur eingeschränkt möglich, aber das ist in der Schule im Rahmen von Klassenarbeiten und ähnlichen Prüfungen nicht unbedingt anders. So lässt sich trotz der eingeschränkten Freiheitsgrade annehmen, dass das Gewähren individueller Freiheiten wie etwa einem selbst bestimmbareren Arbeitstempo oder einer selbst gewählten Sitzhaltung sowie das Gewähren von Nachfragen und ein responsiver Umgang hiermit durchaus förderlich für die Autonomiewahrnehmung eines getesteten Kindes sein dürften. Ebenso sollte das Kompetenzerleben gefördert bzw. unterminiert werden können, et-

wa indem das Kind leistungsunabhängig und wiederholt für seine gute Mitarbeit und Anstrengung gelobt wird.

*Lern- vs. leistungszielorientierte Testsituation.* In verschiedenen theoretischen Strömungen werden zwei grundlegende motivationale Orientierungen unterschieden, die der Bearbeitung von Aufgaben im Lern- und Leistungskontext zugrunde liegen können. Nicholls (1984) stellt die Begriffe Aufgabenorientierung (task involvement) und Ichorientierung (ego involvement) gegenüber. Aufgabenorientierte Personen konzentrieren sich auf die Meisterung der aktuellen Aufgabe ohne sich besondere Sorgen über spätere Bewertungen und Beurteilungen zu machen. Im Gegensatz hierzu fällt es ichorientierten Personen schwer, sich voll und ganz auf die Aufgabe zu fokussieren, da ihre Gedanken fortwährend um die Bewertung ihrer Leistung kreisen, welche sie zudem als indirekte Bewertung ihrer gesamten Person ansehen. Der wahrgenommene Leistungsdruck und die erwarteten negativen Konsequenzen einer schlechten Leistung für das eigene Selbstwertgefühl sind bei ichorientierten Lernern somit deutlich höher ausgeprägt als bei aufgabenorientierten.

Ähnlich der Konzeption von Nicholls (1984) werden in der Forschung zu motivationalen Zielorientierungen Lernziele (mastery goals) von Leistungszielen (performance goals) unterschieden (Dweck & Leggett, 1988; Elliot, 1999; Senko, Hulleman & Harackiewicz, 2011). Während lernzielorientierte Personen bemüht sind, eine konkrete Aufgabe zu meistern und dabei ihre eigenen Kompetenzen zu erweitern, streben leistungszielorientierte Personen danach, bessere Leistungen zu erbringen als andere und ihre Fähigkeiten anderen zu demonstrieren. Aufgrund widersprüchlicher Befunde bezüglich der Effekte der Leistungszielorientierung wurde diese in späteren Arbeiten weiter ausdifferenziert. Gemäß der Erweiterung von Elliot (1999) wollen annäherungs-leistungszielorientierte Personen besser sein als andere und ihre Fähig-

keiten demonstrieren, wohingegen vermeidungs-leistungszielorientierte Lerner bemüht sind, nicht schlechter zu sein als andere und ihre mangelnden Fähigkeiten zu verbergen. In einer Reihe von Studien konnte gezeigt werden, dass eine Annäherungs-Leistungszielorientierung häufig positiv mit pädagogisch-psychologisch bedeutsamen Outcome-Variablen wie z.B. schulischen Leistungen und fachbezogenem Interesse assoziiert ist, wohingegen eine Vermeidungs-Leistungszielorientierung zumeist mit negativen Konsequenzen wie mangelnder Anstrengungsbereitschaft, erhöhter Prüfungsangst oder der Nutzung oberflächlicher Lernstrategien verbunden ist (Senko et al., 2011).

In der Literatur besteht weitgehend Konsens darüber, dass eine durch Instruktionen von Lehrkräften und Testleiterinnen und Testleitern nahegelegte Aufgaben- bzw. Lernzielorientierung positive Effekte und eine external induzierte Ich- bzw. Leistungszielorientierung negative Effekte auf das eigene Selbstwertgefühl sowie die psychologische Anpassung im Allgemeinen hat. Roland (2012) kommt in einer aktuellen Metaanalyse zu den Auswirkungen von Klassenzielstrukturen (d.h. der von der Lehrkraft vermittelten Zielorientierung in der Klasse) zu dem Schluss, dass lernziel- und aufgabenorientierte Instruktionen im Klassenraum positiv mit dem Kompetenzerleben, der Selbstwirksamkeit und dem allgemeinen Selbstwert assoziiert sind. Im Gegensatz hierzu zeigten sich in über 40% der untersuchten Studien negative Effekte einer ich- bzw. leistungszielorientierten Instruktion auf diese Outcomes. Hierbei stellt sich wiederum die Frage, inwiefern solche Befunde aus dem Schulalltag auf individuelle Testsituationen übertragbar sind. Eine Reihe von Studien, in denen eine Lern- vs. Leistungszielorientierung experimentell induziert wurde, liefern jedoch Hinweise auf sehr ähnliche Zusammenhangsmuster. Wie Utman (1997) in seiner Metaanalyse zusammenfasst, waren die Vorteile einer lernzielbasierten Testinstruktion dabei umso deutli-

cher je komplexer das Aufgabenmaterial war. Da das Material bei Intelligenztestungen notwendigerweise komplex sein muss, kann auch für das von uns avisierte Setting einer individuellen Intelligenztestung von einem messbaren Effekt einer lern- vs. leistungszielbetonten Testinstruktion ausgegangen werden.

### **Fragestellung und Hypothesen**

Ergebnisse von Intelligenztestungen einzelner Schülerinnen und Schüler werden in der psychologischen und sonderpädagogischen Praxis sehr häufig als Grundlage für wichtige schulbezogene Entscheidungen herangezogen und sollten daher sowohl valide sein als auch mit möglichst geringen psychischen Belastungen für das jeweilige Kind einhergehen. Ausgehend von früheren Studien, in denen zum einen Effekte unterschiedlichen Testleiterverhaltens auf die Leistung in individuellen Intelligenztestungen (z.B. Petty & Harrell, 1977; Sorrentino et al., 1996) und zum anderen bedeutsame Einflüsse lernzielfokussierter, autonomieunterstützender Test- und Unterrichtssituationen im Allgemeinen berichtet wurden, vermuten wir, dass entsprechend gestaltetes Testleiterverhalten positive bzw. negative Auswirkungen auf den situationalen Selbstwert hat, welche in der Folge auch auf das allgemeine Selbstwertgefühl transferieren. Unseres Wissens existieren zu dieser Frage jedoch bislang keine empirischen Studien. In der vorliegenden Arbeit sollen daher die Effekte unterschiedlich unterstützenden Testleiterverhaltens auf das situationalen Selbstwertgefühl unmittelbar nach einer Intelligenztestung sowie den allgemeinen Selbstwert zwei Wochen nach der Testung untersucht werden. Zur Prüfung dieser Fragestellung werden in einem experimentellen Design zwei verschiedene Experimentalbedingungen (autonomieunterstützendes Testleiterverhalten vs. nicht unterstützendes Testleiterverhalten) und eine Kontrollbedingung hinsichtlich folgender Hypothesen miteinander verglichen.

*Hypothese 1.* Im Vergleich zu der unterstützenden Testbedingung und der Kontrollgruppe sollte sich der situationale Selbstwert der Kinder in der nicht unterstützenden Testbedingung vom ersten zum zweiten Messzeitpunkt verringern.

*Hypothese 2.* Es besteht ein direkter Effekt der Testbedingung auf den situationalen Selbstwert, der nicht vollständig über die Leistung im Intelligenztest mediiert wird.

*Hypothese 3.* Die Veränderungen im situationalen Selbstwert direkt nach der Intelligenztestung transferieren auch auf den allgemeinen Selbstwert zwei Wochen später.

### **Methode**

#### **Stichprobe**

*Rekrutierung der Stichprobe.* Im Rahmen der Rekrutierung wurden sechs Wochen vor Beginn der Sommerferien insgesamt sechs Gymnasien angeschrieben, von denen sich vier damit einverstanden erklärten, dass die Studie in den fünften Klassen im Unterricht kurz vorgestellt und Informationsbriefe an alle interessierten Kinder ausgeteilt wurden. Die Kontaktaufnahme erfolgte über ein Rückmeldeformular, welches interessierte Familien der Testleiterin über die Lehrkräfte zukommen ließen. Teilnahmevoraussetzung war das Einverständnis des Kindes und der Eltern. Als Teilnahmeanreiz wurden den Kindern ein kleines Geschenk direkt nach der Aufgabenbearbeitung sowie die Teilnahme an einer Verlosung von Kino- und Buchgutscheinen in Aussicht gestellt.

*Coverstory.* Da davon ausgegangen wurde, dass das Wissen um die Testung der Intelligenz bei Schülerinnen und Schülern zu erhöhtem Stress geführt hätte, wurde die Rekrutierung unter dem Deckmantel einer Coverstory durchgeführt. Die beteiligten Lehrkräfte, Eltern und Kinder wurden darüber informiert, dass ein neues Förderprogramm zur Unterstützung von Schülerinnen und Schülern mit Lernschwierigkeiten auf

seine Eignung für die fünfte Klassenstufe hin geprüft werden sollte. Bei der Vorstellung der Studie wurden ausdrücklich alle Kinder eingeladen, an der Studie teilzunehmen, unabhängig davon, ob Lernschwierigkeiten bestünden oder nicht. Alle Kinder wurden nach Abschluss der Studie über das tatsächliche Untersuchungsziel und den Grund der Täuschung aufgeklärt. Der gesamte Versuchsablauf und die Coverstory waren zuvor der lokalen Ethikkommission vorgelegt und von dieser als unbedenklich eingeschätzt worden.

*Beschreibung der Stichprobe.* Zu Beginn der Untersuchung bestand die Stichprobe aus 19 Schülerinnen und 13 Schülern ( $N = 32$ ). Eine Versuchsperson entschied sich noch vor der Intelligenztestung nicht weiter an der Untersuchung teilzunehmen. Letztendlich lagen 31 vollständige Datensätze für die weiteren Analysen vor. Zum Zeitpunkt der ersten Erhebung lag das Durchschnittsalter der Kinder bei  $M = 10.97$  ( $SD = 0.32$ ) Jahren. Die Erhebung erstreckte sich insgesamt über einen Zeitraum von zwei Monaten. Aufgrund ihres Testtermins gegen Ende der Sommerferien waren sieben Schülerinnen und Schüler zum Zeitpunkt der letzten Erhebung (Postfragebogen zum allgemeinen Selbstwert) bereits in der sechsten Klasse. Alle anderen Datenerhebungen wurden vor Wiederbeginn der Schule abgeschlossen.

### **Untersuchungsdesign**

*Allgemeiner Ablauf.* Der Ablauf der Studie bestand aus insgesamt vier Messzeitpunkten. Zunächst erhielten alle interessierten Familien zwei Wochen vor Beginn der Sommerferien ein Anschreiben mit Erklärungen zum Ablauf der Studie. Hierin wurden die Kinder gebeten, den beiliegenden ersten Fragebogen auszufüllen und anschließend sofort im Rückumschlag an die Universität zurückzusenden. In diesem Fragebogen wurde das Ausgangsniveau des allgemeinen Selbstwerts vor der Intelligenztestung erfasst (Messzeitpunkt 1). Nach Ein-

gang des Fragebogens wurde mit der jeweiligen Familie ein Termin zur Durchführung der Intelligenztestung vereinbart, der zwei Wochen nach Rücksendung des Fragebogens stattfand. Um bei der Intelligenztestung eine größtmögliche Standardisierung zu gewährleisten, wurden alle Testungen innerhalb der Sommerferien in den Räumlichkeiten der Universität durchgeführt. Um außerdem eine hohe ökonomische Validität sicherzustellen, wurden die Testungen in den Räumen der an den Lehrstuhl für Pädagogische Psychologie angegliederten Beratungsstelle durchgeführt, in denen regelmäßig Sitzungen zur Diagnostik von Lern- und Leistungsstörungen stattfinden und deren Einrichtung mit der einer schulpсихologischen Praxis zu vergleichen ist. Unmittelbar vor und nach der Testung wurden die Kinder gebeten, jeweils einen Fragebogen zu ihrem situationalen Selbstwert auszufüllen (Messzeitpunkt 2 und 3). Abschließend erhielten alle Kinder zwei Wochen nach der Testung postalisch einen letzten Fragebogen, in dem der allgemeine Selbstwert nach der Intelligenztestung erfasst wurde (Messzeitpunkt 4).

*Zuweisung zu den Testbedingungen.* Um die erwarteten Effekte einer unterstützenden und nicht-unterstützenden Durchführung des Intelligenztests prüfen zu können, wurden zu Beginn der Studie alle Probanden randomisiert einer von drei Testbedingungen zugewiesen. Insgesamt befanden sich elf Kinder (vier Jungen, sieben Mädchen) in der nicht-unterstützenden Bedingung (EG 1), während die Gruppen der unterstützenden Bedingung (EG 2, vier Jungen, sechs Mädchen) und die Kontrollgruppe mit Alternativaufgabe (fünf Jungen, fünf Mädchen) jeweils aus zehn Kindern bestanden. Die Testungen wurden jeweils als Einzeltestungen durchgeführt. Eine standardisierte, eng am Manual orientierte Durchführung des Tests wurde durch eine ausführliche Schulung der Testleiterinnen vor Beginn der Studie sichergestellt.

*Nicht-unterstützende Testbedingung (EG 1).* In der nicht-unterstützenden Bedin-

gung wurde die Interaktion zwischen der Testleiterin und dem Kind sowie der Begleitperson so gering wie möglich gehalten. Um den offiziellen Eindruck zu verstärken, erschien die Testleiterin zur Begrüßung mit dem Testkoffer in der Hand. Das Kind und die Begleitperson wurden knapp und förmlich begrüßt. Anschließend wurde das Kind gebeten, den Fragebogen zum situationalen Selbstwert selbstständig auszufüllen. Bei Verständnisfragen wendeten sich die Kinder an ihre Begleitperson. Nach Ausfüllen des Fragebogens forderte die Testleiterin das Kind auf, ihr zu folgen. Die Begleitperson wurde gebeten, auf dem Flur zu warten. Im Versuchsraum wurde dem Kind ein Stuhl zugewiesen. Nachdem das Testmaterial vor dem Kind aufgebaut worden war, verblieb der Testkoffer auf dem Tisch. Vor Beginn der Aufgabenbearbeitung wurde dem Kind erklärt, dass ihm Testaufgaben vorgelegt würden, die es lösen solle. Dabei sei es besonders wichtig, dass es möglichst viele Testaufgaben richtig löse und dass später überprüft würde, wie gut ihm dies gelungen sei. Zusätzlich wurde es aufgefordert sich Mühe zu geben, damit es so viele Testaufgaben wie möglich lösen könne. Während der gesamten Testsituation wurde dem Kind kein Feedback erteilt, sowie keine Verstärkung durch freundliches Verhalten gegeben. Nach der Durchführung des Intelligenztests wurde dem Kind erneut der Fragebogen zum situationalen Selbstwert vorgelegt. Im Anschluss konnte das Kind an der Verlosung teilnehmen und sich dann ein Geschenk auswählen, wobei die Testleiterin auch hier ihre distanzierte Haltung gegenüber dem Kind beibehielt. Zuletzt wurde das Kind zu seiner Begleitperson zurückgebracht. Abschließend wurde darauf hingewiesen, dass ca. zwei Wochen nach der Testung erneut ein Fragebogen zugeschickt werden würde, mit der Bitte, diesen schnellstmöglich zurückzusenden, da nur Kinder, von denen vollständige Daten vor-

lägen, an der Verlosung teilnehmen könnten.

*Unterstützende Testbedingung (EG 2).* In der unterstützenden Testbedingung sollte sich der gesamte Ablauf für die Kinder so stressfrei wie möglich gestalten und so wenig Aufregung wie möglich verursachen. Daher wurde das Testmaterial bereits vor der Begrüßung im Testraum aufgebaut. Zu Beginn begrüßte die Testleiterin das Kind und seine Begleitperson freundlich und bat das Kind, zunächst den Fragebogen zum situationalen Selbstwert auszufüllen. Anschließend durfte das Kind gemeinsam mit seiner Begleitperson den Raum ansehen, in dem die Aufgabenbearbeitung stattfinden würde. Es folgte ein kurzes Gespräch, um dem Kind die Gelegenheit zu geben, sich im Beisein der Begleitperson mit Situation und Umgebung vertraut zu machen. Indem das Gespräch erst nach Ausfüllen des Fragebogens stattfand, sollte gewährleistet werden, dass die Kinder in allen Gruppen den ersten Fragebogen unter ähnlichen Bedingungen ausfüllten und sich das Gespräch nicht auf das Antwortverhalten der zweiten Gruppe auswirkte. Anschließend wurde die Begleitperson gebeten, den Raum zu verlassen. Vor Beginn der Aufgabenbearbeitung wurde dem Kind erklärt, dass man sich nun zusammen Rätsel anschauen würde, die bereits viele Kinder bearbeitet und die bisher allen viel Spaß bereitet hätten<sup>1</sup>. Weiterhin wurde erläutert, dass die Aufgaben mit der Zeit schwieriger würden, es aber nicht schlimm sei, wenn das Kind diese nicht lösen könne, weil solche Aufgaben auch für ältere Kinder seien. Während der Testung wurde dem Kind leistungsunabhängig positives Feedback gegeben, indem es wiederholt für seine gute Mitarbeit und Anstrengung gelobt wurde. Nach der Testung wurde ein vom Kind selbst gewähltes Spiel gespielt. Der Abschluss erfolgte analog zur nicht unterstützenden Testbedingung mit dem Unterschied, dass das Kind nach Aus-

<sup>1</sup> Diese Aussagen werden im Manual der K-ABC als mögliche Hinweise genannt, die dem Kind gegeben werden dürfen, und weichen somit nicht von der standardisierten Instruktion ab.



füllen des Fragebogens und abschließend noch einmal vor der Begleitperson für seine gute Mitarbeit gelobt wurde.

**Kontrollgruppe.** Die Kinder der Kontrollgruppe füllten ebenfalls zunächst den Fragebogen zum situationalen Selbstwert aus. Anschließend wurden sie gebeten, der Testleiterin in den Testraum zu folgen, während die Begleitperson auf dem Flur vor dem Testraum warten sollte. In der Kontrollgruppe wurde kein Intelligenztest durchgeführt, um neben der Prüfung der postulierten Hypothesen auch einschätzen zu können, ob aus der Durchführung eines Intelligenztests unter unterstützenden Bedingungen keine stärkere affektive Belastung resultiert als aus der Durchführung einer für den schulischen Unterricht typischen Beurteilungsaufgabe. Dementsprechend wurde das Kind gebeten, drei kurze narrative Texte, die für die Studie aus altersgerechten Büchern zusammengestellt wurden, zu lesen und hierzu jeweils drei Fragen zu beantworten. Die Aufgaben wurden so konzipiert, dass ihre Bearbeitung ungefähr genauso lange dauerte wie die Intelligenztestung in den anderen Bedingungen. Das Kind wurde darauf hingewiesen, dass es keine falschen Antworten gebe, da nur die Meinung des Kindes wichtig sei, und dass auch kein Zeitdruck bestehe. Während der Bearbeitung der Texte blieb die Testleiterin mit am Tisch, es fand jedoch keine Interaktion mit dem Kind statt. Der abschließende Teil gestaltete sich analog zur unterstützenden Testbedingung.

### **Untersuchungsverfahren**

**Kaufmann-Assessment Battery for Children.** Zur Testung der Intelligenz wurde die Kaufmann-Assessment Battery for Children (K-ABC; Melchers & Preuss, 2009) eingesetzt. Die K-ABC ist ein standardisierter und in der Praxis häufig eingesetzter Test zur Erfassung der intellektuellen Leistungsfähigkeit von Kindern im Altersbereich von 2;5 bis 12;5 Jahren. Für den Zweck der vorliegenden Studie war die Verwendung der K-ABC aus zwei Gründen besonders geeignet: Einerseits lagen die Kinder mit durchschnittlich ca. elf Jahren an der oberen Grenze des Altersbereichs des Testverfahrens, wodurch die Möglichkeit einer wiederholten Anwendung der K-ABC als Intelligenzdiagnostikum in der weiteren Entwicklung der Kinder nahezu ausgeschlossen werden konnte. Zweitens wurde angenommen, dass eine gewisse Anzahl von Aufgaben für die Kinder vergleichsweise einfach zu lösen waren, sodass mögliche Veränderungen im Selbstwert nach der Testung tatsächlich auf das Verhalten der Testleiterin zurückzuführen sein dürften und nicht auf eine Überforderung durch die Testaufgaben.

**Fragebogen zum allgemeinen Selbstwert.** Der allgemeine Selbstwert wurde mit einer Skala von Rosenberg in der deutschen Übersetzung von Ferring und Filipp (1996) erfasst<sup>2</sup>. Die Rosenberg-Skala hat sich als Standardinstrument zur Selbstwert-Messung bewährt. Sie besteht aus 10 Items (z.B. „Ich glaube, ich habe eine Menge guter Eigenschaften.“) und misst eindimensional die Höhe des allgemeinen Selbstwerts. Alle Items waren auf einer Ratingskala von 1 (*stimmt gar nicht*) bis 5 (*stimmt genau*) zu

<sup>2</sup> Die deutsche Fassung von Ferring und Filipp (1996) wurde von Collani und Yorck Herzbergs (2003) Version vorgezogen. Auf der einen Seite unterscheiden sich die beiden Fassungen nur durch eine geringfügige Formulierungsänderung des vierten Items. Auf der anderen Seite wurde in der therapeutischen und sonderpädagogischen Praxis und Forschung die Fassung von Ferring und Filipp (1996) bei Weitem häufiger eingesetzt, was eine präzisere Vergleichbarkeit mit den hier berichteten Befunden erlaubt. In der Abwägung beider Punkte schien uns der letztgenannte der wichtigere Aspekt zu sein.

beantworten<sup>3</sup>. Die interne Konsistenz der Skala erwies sich sowohl vor ( $\alpha = .71$ ) als auch nach der Testung ( $\alpha = .84$ ) als zufriedenstellend.

*Fragebogen zum situationalen Selbstwert.* Um zu erfassen, wie sich die teilnehmenden Kinder unmittelbar vor und nach der Testsituation selbst bewerten, wurde eine modifizierte Version der State Self-Esteem Scale (Heatherton & Polivy, 1991; Schwinger, 2010) verwendet. Anhand dieser Skala können kurzfristige Veränderungen im Selbstwert einer Person zuverlässig erfasst werden. Die ursprüngliche Version beinhaltet Items zum situationalen leistungsbezogenen, sozialen und äußerlichen Selbstwert. Für die adaptierte Fassung wurden vier Items zum leistungsbezogenen und sozialen Selbstwert ausgewählt, da durch unterschiedliches Testleiterverhalten und Intelligenztestungen diese beiden Selbstwertbereiche am stärksten angesprochen werden sollten. Die Items wurden an den Kontext der Testsituation angepasst und durch zwei weitere selbst entwickelte Items ergänzt, sodass die Skala insgesamt aus sechs Items bestand (z.B. „Im Moment glaube ich, dass ich nichts richtig kann.“). In früheren Untersuchungen hatte sich diese adaptierte Skala als eindimensional (vgl. auch die Option eines Globalfaktors bei Heatherton und Polivy, 1991), reliabel und angemessen veränderungssensitiv erwiesen, auch im Vergleich zur Rosenberg-Skala (Schwinger, 2010). Wie bei der Rosenberg-Skala wurden die Items auf einer Likertskala von 1 (*stimmt gar nicht*) bis 5 (*stimmt genau*) beantwortet. Die internen Konsistenzen waren ebenfalls akzeptabel ( $\alpha = .60/.67$ ).

## Ergebnisse

### Ausgangsbedingungen

Um zu prüfen, ob sich die Kinder in den verschiedenen Testbedingungen hinsichtlich ihres Ausgangsniveaus in den untersuchten Variablen unterscheiden, wurden einfaktorielle Varianzanalysen durchgeführt. Diese Analysen ergaben keine signifikanten Gruppenunterschiede in Bezug auf den allgemeinen Selbstwert ( $F [2,27] = 0.23, p = .80$ ) und situationalen Selbstwert ( $F [2,28] = 1.03, p = .37$ ) vor der Testung sowie in Bezug auf den ermittelten IQ-Wert ( $F [1,19] = 0.09, p = .77$ ). Die Mittelwerte und Standardabweichungen aller Skalen sind Tabelle 1 zu entnehmen.

### Effekte der Intelligenztestung

Trotz der ähnlichen Ausgangsbedingungen wurden die Effekte der Testbedingung auf den situationalen Selbstwert durch eine Kovarianzanalyse untersucht, in die der Prä-Wert im situationalen Selbstwert sowie das Geschlecht<sup>4</sup> und der allgemeine Selbstwert vor der Testung als Kovariaten aufgenommen wurden. Während das Geschlecht und der allgemeine Selbstwert keinen signifikanten Einfluss auf den situationalen Selbstwert nach der Testung hatten, trug der situationale Selbstwert vor der Testung signifikant zur Erklärung des situationalen Selbstwerts nach der Testung bei ( $F [1,24] = 9.47, p < .01$ ). Nach Ermittlung der auf die Kovariaten entfallenen Varianzanteile konnten die verbleibenden Unterschiede im situationalen Selbstwert nach der Testung zudem

<sup>3</sup> Die ursprüngliche Fassung der Rosenberg-Skala wurde mit vier Antwortstufen konzipiert. Das fünfstufige Antwortformat wurde hier aus zwei Gründen gewählt. Zum einen bietet eine mittlere Antwortkategorie den Kindern die Möglichkeit, eine evtl. indifferente Meinung zu einzelnen Fragen auszudrücken (siehe Rost, 2004, für eine detaillierte Diskussion). Zum anderen wurde das fünfstufige Antwortformat bei der State Self-Esteem Scale von Heatherton und Polivy (1991) verwendet. Um die Kinder nicht zu verwirren und die ökologische Validität der Antworten nicht unnötig zu gefährden, erschien es sinnvoll, für die Selbstwertmessungen die gleichen Antwortformate einzusetzen.

<sup>4</sup> Eine Reihe von Studien berichten Geschlechtsunterschiede im Selbstkonzept (z.B. Tiedemann & Faber, 1995) und Selbstwert (z.B. Kling, Hyde, Showers & Buswell, 1999). Daher erschien es uns sinnvoll, den Faktor Geschlecht als mögliche Störvariable in den Analysen zu kontrollieren.

Tabelle 1: Mittelwerte und Standardabweichungen in den verschiedenen Versuchsbedingungen

	Gesamt	EG 1	EG 2	KG
	M (SD)	M (SD)	M (SD)	M (SD)
Situationaler Selbstwert vorher	4.07 (0.51)	4.12 (0.61)	4.20 (0.41)	3.88 (0.48)
Situationaler Selbstwert nachher	4.15 (0.57)	3.83 (0.64)	4.30 (0.37)	4.35 (0.55)
Allgemeiner Selbstwert vorher	3.98 (0.51)	4.06 (0.59)	3.90 (0.43)	3.98 (0.55)
Allgemeiner Selbstwert nachher	4.21 (0.57)	4.12 (0.77)	4.25 (0.47)	4.25 (0.49)
Intelligenz (IQ)	100.48 (9.48)	101.09 (9.43)	99.80 (9.98)	—

Anmerkungen. EG 1 = nicht-unterstützende Testbedingung; EG 2 = unterstützende Testbedingung; KG = Kontrollgruppe.

durch die Gruppenzugehörigkeit erklärt werden (Faktor Gruppe:  $F [1,24] = 5.85$ ,  $p < .01$ ). Die Effektstärken lagen hierbei sowohl für den Einfluss des situationalen Selbstwerts vor der Testung ( $\eta^2 = .28$ ) als auch für den Einfluss der Gruppenzugehörigkeit ( $\eta^2 = .33$ ) im hohen Bereich. Paarweise Vergleiche der drei Gruppen unter Kontrolle der Kovariaten sowie mit Adjustierung des Alpha-Fehlers zeigten im Einklang mit Hypothese 1, dass der Effekt der Gruppenzugehörigkeit darauf zurückzuführen ist, dass die Kinder in der nicht-unterstützenden Bedingung nach dem Intelligenztest signifikant geringere Werte im situationalen Selbstwert aufwiesen als die Kinder in der unterstützenden Bedingung (Differenz:  $M = 0.38$ ,  $SD = 0.16$ ;  $p < .05$ ) und die Kinder in der Kontrollgruppe (Differenz:  $M = 0.58$ ,  $SD = 0.17$ ;  $p < .05$ ). Die Werte der Kinder in der unterstützenden Bedingung und in der Kontrollgruppe unterschieden sich dagegen nicht signifikant.

Im Anschluss wurden zur Prüfung der angenommenen Transfer- und Mediationseffekte (Hypothesen 2 und 3) Partial Least Squares (PLS) Pfadanalysen mit dem Programm WarpPLS 4.0 (Kock, 2013) berechnet. PLS Regressionsmodelle stellen eine Erweiterung der bekannteren linearen Regressionsmodelle dar und zeichnen sich vor allem durch eine höhere Flexibilität bei der Modellspezifikation aus. Besonders hervorzuheben sind in diesem Zusammenhang a) die Robustheit von PLS-Modellen gegen-

über Verletzungen der Normalverteilung einzelner Modellvariablen sowie b) die Möglichkeit komplexe Modelle auch bei sehr kleinen Stichproben valide und reliabel berechnen zu können (Kock, 2013). In dem hier spezifizierten Pfadmodell wurde die Höhe des allgemeinen Selbstwerts nach der Intelligenztestung durch die Gruppenzugehörigkeit vorhergesagt, wobei der situationalen Selbstwert als Mediator modelliert wurde und wiederum - wie in den vorherigen Analysen - das Ausgangsniveau der Variablen kontrolliert wurde (siehe Abbildung 1). Um zu prüfen, inwiefern die Gruppenzugehörigkeit den situationalen Selbstwert auch unabhängig von der Leistung im Intelligenztest beeinflusst, wurde in dem Pfadmodell sowohl ein direkter Pfad der Gruppenzugehörigkeit auf den situationalen Selbstwert als auch ein indirekter Pfad über die Leistung im Intelligenztest ergänzt. Da mit den Kindern der Kontrollgruppe kein Intelligenztest durchgeführt wurde und somit keine Leistungsdaten vorlagen, wurden nur die Kinder der unterstützenden und der nicht-unterstützenden Bedingungen ( $n = 21$ ) in die Analysen einbezogen.

Insgesamt klärten die Prädiktoren 81 % der Varianz des situationalen Selbstwerts sowie 78 % der Varianz des allgemeinen Selbstwerts nach der Intelligenztestung auf. Im Einklang mit den Hypothesen zeigte sich unter Kontrolle der Stabilitäten nach der Durchführung der Intelligenztestung ein signifikanter Effekt der Gruppenzugehörigkeit

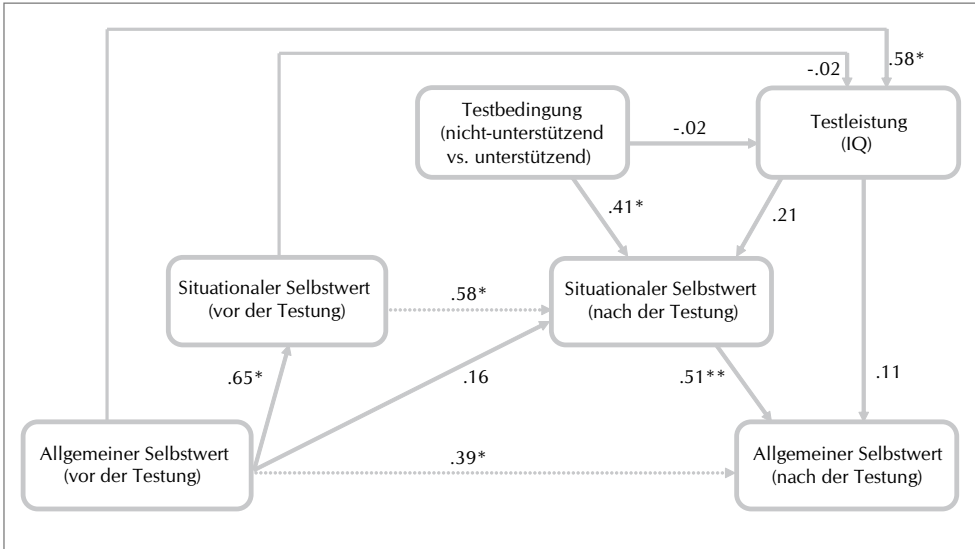


Abbildung 1: Pfadmodell zum Einfluss der Testbedingung auf die Testleistung, den situationalen und den allgemeinen Selbstwert ( $\chi^2 [df = 2, N = 21] = 0.800; p = .93; CFI = 1.00; RMSEA = .00$ ).

Anmerkungen. Gestrichelte Pfeile symbolisieren Stabilitätskoeffizienten. \*  $p < .05$ , \*\*  $p < .01$ . Die Testbedingungen wurden folgendermaßen kodiert: nicht-unterstützend = 0; unterstützend = 1.

auf den situationalen Selbstwert ( $\beta = .41$ ,  $p < .01$ ) sowie ein signifikanter Effekt des situationalen Selbstwerts auf den allgemeinen Selbstwert ( $\beta = .51$ ,  $p < .01$ ). Die Leistung im Intelligenztest zeigte keinen direkten Einfluss auf den situationalen Selbstwert ( $\beta = .21$ ) und war zudem unabhängig von der Gruppenzugehörigkeit ( $\beta = -.02$ ).

## Diskussion

Ziel der vorliegenden Studie war es zu überprüfen, inwieweit unterstützendes vs. nicht unterstützendes Verhalten von Testleiterinnen und Testleitern im Kontext von Intelligenztestungen das Selbstwertgefühl von Kindern beeinflusst. Darüber hinaus wurde untersucht, inwiefern dieser Effekt über die im Intelligenztest erbrachte Leistung vermittelt ist oder sich direkt auf das Selbstwertgefühl der Kinder auswirkt. Weiterhin wurde der Frage nachgegangen, ob die Effekte nicht nur kurzfristig, d.h. direkt im Anschluss an die Testung, auftreten, sondern

auch noch mittelfristig, d.h. zwei Wochen nach der Testung, nachweisbar sind. Um diese Hypothesen zu prüfen, wurden zwei Experimentalgruppen mit unterstützenden vs. nicht unterstützenden Instruktionen sowie eine Kontrollgruppe miteinander verglichen.

In Einklang mit unserer ersten Hypothese zeigten die Kinder der nicht unterstützenden Testbedingung nach der Testung signifikant niedrigere Werte im situationalen Selbstwert als die Kinder der anderen beiden Gruppen. Eine Stärke der vorliegenden Studie im Vergleich zu früheren Untersuchungen (z.B. Guay et al., 2008) besteht darin, dass der Effekt auch unter Kontrolle des Geschlechts sowie des allgemeinen und situationalen Selbstwerts vor der Testung abgesichert werden konnte. Zum einen bestätigen unsere Ergebnisse somit die in der Literatur berichteten Befunde zu bedeutsamen Auswirkungen verschiedener Testinstruktionen. Zum anderen erweitert unsere Studie die empirische Evidenz in diesem Zusammenhang, da in der vorliegenden

Untersuchung nicht auf die Testangst, sondern auf den situationalen Selbstwert fokussiert wurde, Kinder anstatt von Erwachsenen untersucht wurden und zur Absicherung der Befunde eine Kontrollgruppe ohne Intelligenztestung miteinbezogen wurde.

Die Ergebnisse der Pfadanalyse stützten auch unsere zweite Hypothese, wonach ein nicht unterstützendes Testleiterverhalten einen direkten negativen Einfluss auf das Selbstwertgefühl haben sollte, der nicht vollständig über die erzielte Leistung im Intelligenztest vermittelt wird. Unserem Wissen nach wurde dieser direkte Link zwischen Testinstruktion und Selbstwertgefühl bislang nicht explizit untersucht. Wir stimmen jedoch mit der Ansicht von Bless und Mohr (2007) überein, dass die Herstellung eines gesunden Selbstwertgefühls und eines hohen psychischen Wohlbefindens wichtige Ziele schulischer Förderung darstellen, die neben dem Streben nach guten schulischen Leistungen nicht aus dem Blick geraten dürfen. Diesbezüglich verdeutlichen die pfadanalytischen Befunde umso mehr die Wichtigkeit eines autonomieförderlichen und lernzielorientierten Instruktionsstils für das kindliche Wohlbefinden. Die Bedeutsamkeit der direkten Auswirkung des Testleiterverhaltens wird außerdem durch die Tatsache untermauert, dass in unserer Stichprobe kein Effekt der Testbedingung auf die Leistung in der K-ABC festzustellen war und darüber hinaus der indirekte Effekt auf den situationalen Selbstwert nach der Testung nicht signifikant war. Somit belegen unsere Befunde einen Effekt des Testleiterverhaltens, der unabhängig von der erzielten Leistung auftreten kann und der wahrscheinlich gerade deshalb in Forschung und Praxis wenig Beachtung gefunden hat. Das Ausbleiben eines signifikanten Effekts der Testbedingung auf die Testleistung steht zwar im Widerspruch zu den Ergebnissen anderer Autoren (z.B. Meijer & Oostdam, 2007; Sorrentino et al., 1996), lässt sich unserer Meinung nach aber plausibel durch die hier gewählte Kombination von Stichprobenalter und Testverfahren erklären. Da die K-ABC

nur bis zum Alter von 12 Jahren normiert ist, konnten die hier getesteten Kinder eine ausreichende Anzahl von Aufgaben maßstäblich relativ leicht lösen. Auch die in Tabelle 1 dargestellte Standardabweichung verdeutlicht, dass die Varianz in den Testleistungen generell eingeschränkt war, was einen möglichen signifikanten Einfluss der Testinstruktion auf die Testleistung aus methodischen Gründen unwahrscheinlicher macht.

Ein weiterer Aspekt, der in der früheren Forschung bislang ebenfalls wenig beachtet wurde, bezieht sich auf mögliche Transfereffekte eines bestimmten Verhaltens in der Testsituation in den kindlichen Alltag einige Wochen später. Unsere pfadanalytische Auswertung legt nahe, dass sich die Effekte der verschiedenen Testinstruktionen über den situationalen Selbstwert hinaus auch im allgemeinen Selbstwert zwei Wochen später niedergeschlagen haben. Allerdings bleibt die Interpretation dieses Befundes aus zwei Gründen eher spekulativ. Erstens haben wir den allgemeinen Selbstwert nach der Testung nur ein weiteres Mal (zwei Wochen nach der Testung) erhoben. Somit können keine Aussagen über eine detailliertere, weitergehende Entwicklung des Selbstwerts getroffen werden. Zweitens wurden die Kinder und Eltern nicht bezüglich potentiell selbstwertrelevanter Ereignisse in den Tagen nach der Testung befragt. Der hier gefundene Transfereffekt auf den allgemeinen Selbstwert sollte daher in zukünftigen Studien wieder aufgegriffen und genauer analysiert werden. Für den geeigneten pädagogischen Umgang mit Kindern, die negative Testerfahrungen gemacht haben, dürfte ein besseres Verständnis derartiger mittel- und längerfristiger Effekte von zentraler Bedeutung sein.

### *Praktische Implikationen*

Die Ergebnisse der Studie leisten eine Reihe wertvoller Implikationen für die Praxis. Intelligenztestungen stellen ein bewährtes und grundlegendes Instrument zur Abklä-

rung verschiedener Probleme im Kindes- und Jugendalter dar, wenn es beispielsweise um die Diagnostik von Lern- und Leistungsstörungen, die Feststellung eines sonderpädagogischen Förderbedarfs oder um die Beurteilung (komorbider) psychischer Störungen geht. Das Durchführen von Intelligenztests gehört somit in vielen schulpсихologischen, psychotherapeutischen, sonderpädagogischen und auch kinderpsychiatrischen Einrichtungen zum beruflichen Alltag. Die Ergebnisse der vorliegenden Studie unterstreichen, welche Bedeutung der jeweiligen Durchführung eines Intelligenztests hierbei nicht nur in Bezug auf eine valide Erfassung der Leistung, sondern auch im Hinblick auf das Wohlbefinden eines Kindes zukommt.

Hierbei ist zu betonen, dass sich die Kinder in der unterstützenden Testbedingung nach der Testung in ihrem Selbstwertgefühl nicht von den Kindern in der Kontrollgruppe unterschieden. Diese Ergebnisse weisen darauf hin, dass eine Intelligenztestung für Kinder grundsätzlich keine größere affektive Belastung darstellt als die Bearbeitung einer typischen Schulaufgabe. Eine zentrale Voraussetzung scheint hierbei jedoch zu sein, dass die Art und Weise, in der die Testung durchgeführt wird, das Autonomie- und Kompetenzerleben des Kindes nicht untergräbt, indem beispielsweise eine aufgabenorientierte Bearbeitung ermöglicht wird und positives anstrengungsbezogenes Feedback gegeben wird. Umgekehrt verdeutlichen die Ergebnisse aber auch, dass eine Verletzung dieser Bedürfnisse auch in einer einmaligen, klar umgrenzten und vorübergehenden Situation wie einer Intelligenztestung deutliche affektive Konsequenzen für ein Kind haben kann. Bei der Interpretation der Befunde ist zu berücksichtigen, dass die Testleiterinnen in der nicht unterstützenden Bedingung sich weder unfreundlich verhielten noch negative Kommentare äußerten, sondern lediglich dem Kind gegenüber eine neutrale Haltung einnahmen, allgemein auf Feedback verzichteten und im Rahmen der Instruktion beim Kind eine Ich- bzw. Leis-

tungszielorientierung induzierten. Vor dem Hintergrund, dass die Durchführung von Intelligenztests eine Routineuntersuchung darstellt, die in der Regel unter hohem Zeitdruck durchgeführt werden muss, ist zu erwarten, dass ein solches Verhalten in der Praxis durchaus vorkommen kann. Darüber hinaus werden Eltern in der Regel im Vorfeld einer Intelligenztestung nicht darüber informiert, wie sie ihre Kinder auf die anstehende Testung vorbereiten sollen. Dementsprechend sollte eine mit der nicht unterstützenden Bedingung vergleichbare Testsituation auch bereits dann entstehen, wenn ein Kind im Vorfeld von seinen Eltern erfährt, dass seine Intelligenz „getestet“ werden wird, und anschließend auf eine Testleiterin oder einen Testleiter trifft, die oder der sich eher distanziert verhält und auf Grund des vorherrschenden Zeitdrucks auf eine kurze Eingewöhnungsphase vor der Testung, auf anstrengungsbezogenes Feedback oder auf einen positiven Abschluss der Situation verzichtet. Insbesondere das Ergebnis, dass eine nicht unterstützende Testsituation auch zwei Wochen nach der Testung noch Auswirkungen auf den allgemeinen Selbstwert hatte, zeigt, wie wichtig es ist, nicht nur praktizierende und angehende Diagnostikerinnen und Diagnostiker, sondern auch Eltern und Lehrkräfte für einen angemessenen und kindgerechten Umgang mit einer Testsituation zu sensibilisieren und ihnen zu verdeutlichen, welche Risiken mit der Durchführung einer Intelligenztestung potentiell verbunden sein können.

### *Methodische Limitationen*

Auch wenn die Studie wertvolle praktische Implikationen leistet, unterliegt sie einigen Limitationen. Eine zentrale Einschränkung bezieht sich auf die kleine Stichprobe und die damit verbundene Unsicherheit bezüglich der Generalisierbarkeit unserer Befunde. Zwar lagen die Effektstärken durchweg in einem recht hohen Bereich und die zentralen Zusammenhänge waren auch signifikant. Dennoch handelte es sich bei den Kin-

dern in der vorliegenden Studie ausschließlich um Schülerinnen und Schüler, die die fünfte Klasse eines Gymnasiums besuchten. Aus zwei Gründen wurde bei der Konzeption der Studie bewusst die Entscheidung getroffen, die Studie mit Gymnasiasten durchzuführen. Erstens wurde davon ausgegangen, dass diese Kinder durchschnittlich recht gut mit der Bearbeitung der Testaufgaben zurecht kommen sollten, insbesondere da sie bereits die obere Altersgrenze des Intelligenztests erreicht hatten. Hierdurch sollte die allgemeine Belastung durch die Teilnahme an der Studie so gering wie möglich gehalten werden. Zweitens erwarteten wir, dass - sofern die Effekte bereits in dieser insgesamt eher unbelasteten Stichprobe nachweisbar sind - vergleichbare oder sogar deutlichere Effekte auch bei Kindern, die durch schulische Misserfolgserlebnisse bereits vorbelastet sind, auftreten sollten. Trotzdem können die Ergebnisse nicht ohne Weiteres auf Schülerinnen und Schüler anderer Schulformen generalisiert werden. So ließe sich etwa vermuten, dass Gymnasiasten ihren Selbstwert generell stärker von schulischen Leistungen abhängig machen als Kinder in leistungsschwächeren Schulsettings. Weitere Aspekte, die in zukünftigen Studien berücksichtigt werden sollten, um die Aussagekraft und Generalisierbarkeit der Befunde zu erhöhen, beziehen sich auf die Untersuchung anderer Altersgruppen sowie emotional stärker vorbelasteter Kinder, die Vermeidung von Selbstselektionseffekten bei der Stichprobenerhebung, die Variation der verwendeten Intelligenztestverfahren und die Kontrolle der Erhebungssituation beim Ausfüllen der Fragebögen.

### Ausblick

Die vorliegende Studie verdeutlicht, dass die Durchführung einer Intelligenztestung für Kinder nicht emotional belastend sein muss, es aber unter bestimmten Bedingungen durchaus sein kann. In dem experimentellen Design konnte unter bestmöglichem

Ausschluss von Störfaktoren gezeigt werden, wie wichtig es ist, während der Testung eine Atmosphäre zu schaffen, die eine lernzielorientierte Bearbeitung der Aufgaben ermöglicht und damit dem Autonomie- und Kompetenzbedürfnis des Kindes entspricht. Nur wenn diese Voraussetzungen erfüllt sind, kann gewährleistet werden, dass die Intelligenz eines Kindes valide erfasst wird, ohne dass die Durchführung für das Kind mit emotionalen Kosten verbunden ist. Trotz der offen gebliebenen Fragen glauben wir, dass dieses Ergebnis von zentraler Bedeutung für die praktische Arbeit aller Berufsgruppen ist, die sich mit der Diagnostik der Intelligenz im Kindes- und Jugendalter beschäftigen, und eine Grundlage für weitergehende Forschung bilden kann.

### Literaturverzeichnis

- Bless, G. & Mohr, K. (2007). Die Effekte von Sonderschulunterricht und Gemeinsamen Unterricht auf die Entwicklung von Kindern mit Lernbehinderungen. In: J. Walter & F. Wember (Hrsg.). *Sonderpädagogik des Lernens* (S. 385-392). Stuttgart: Hogrefe-Verlag.
- Collani, G. & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 3-7.
- Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. Cambridge University Press.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum
- Dweck, C. S. & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34, 169-189.

- Ferring, D. & Filipp, S.-H. (1996). Messung des Selbstwertgefühls: Befunde zu Reliabilität, Validität und Stabilität der Rosenberg-Skala. *Diagnostica, 42*, 284-292.
- Fischbach, A., Schuchardt, K., Mähler, C. & Hasselhorn, M. (2010). Zeigen Kinder mit schulischen Minderleistungen sozio-emotionale Auffälligkeiten? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 42*, 201-210.
- Greenberg, J. & Pyszczynski, T. (1985). Compensatory self-inflation: A response to the threat to self-regard of public failure. *Journal of Personality and Social Psychology, 49*, 273-280.
- Guay, F., Delisle, M. N. & Fernet, C. (2008). Does task-related identified regulation moderate the sociometer effect? A study of performance feedback, perceived inclusion, and state self-esteem. *Social Behavior and Personality: an International Journal, 36*, 239-254.
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of 800 + metaanalyses on achievement*. Oxford, England: Routledge.
- Heatherton, T. F. & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*, 895-910.
- Heisler, G. & Schill, T. R. (1972). An investigation of ways to induce optimal aptitude-performance. *Proceedings of the 80th Annual Convention of the American Psychological Association, 7*, 483-484.
- Jussim, L. & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review, 9*, 131-155.
- Kling, K. C., Hyde, J. S., Showers, C. J. & Buswell, B. N. (1999). Gender differences in self-esteem: a meta-analysis. *Psychological Bulletin, 125*, 470-500.
- Kock, N. (2013). *WarpPLS 4.0 User Manual*. Laredo, Texas: ScriptWarp Systems
- Marsh, H. W. & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*, 133-163.
- Meijer, J. & Oostdam, R. (2007). Test anxiety and intelligence testing: A closer examination of the stage-fright hypothesis and the influence of stressful instruction. *Anxiety, Stress and Coping, 20*, 77-91.
- Meijer, J. & Oostdam, R. (2011). Effects of instruction and stage-fright on intelligence testing. *European Journal of Psychology of Education, 26*, 143-161.
- Melchers, P. & Preuß, U. (2009). *Kaufman-Assessment Battery for Children, Deutsche Version (K-ABC; 8., unveränd. Aufl.)*. Frankfurt am Main: Pearson Assessment.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice and performance. *Psychological Review, 91*, 328-346.
- Petty, E. N. & Harrell, E. H. (1977). Effect of programmed instruction related to motivation, anxiety and test wiseness on group IQ test performance. *Journal of Educational Psychology, 69*, 630-635.
- Reeve, J. & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology, 98*, 209-218.
- Rolland, R. G. (2012). Synthesizing the evidence on classroom goal structures in middle and secondary schools: A meta-analysis and narrative review. *Review of Educational Research, 82*, 396-435.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78.
- Sarason, I. G. (1961). The effects of anxiety and threat on the solution of a difficult task. *Journal of Abnormal and Social Psychology, 62*, 165-168.



- Sassenrath, J. M. & Kight, H. R. (1965). Anxiety, anxiety reduction and motivating instructions in human learning and performance. *Psychological Reports*, 16, 243-250.
- Schwinger, M. (2010). *Die Wirksamkeit von Self-Handicapping als Strategie zur Selbstwertregulation*. Vortrag gehalten auf dem 47. Kongress der Deutschen Gesellschaft für Psychologie, Bremen.
- Senko, C., Hulleman, C. S. & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist*, 46, 26–47.
- Sorrentino, R. M., Brouwers, M. C., Hanna, S. E. & Roney, C. R. (1996). The nature of the test taking situation: Informational and affective influences on intelligence test performance. *Learning and Individual Differences*, 8, 105-120.
- Tiedemann, J. & Faber, G. (1995). Mädchen im Mathematikunterricht: Selbstkonzept und Kausalattributionen im Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27, 61-71.
- Utman, C. H. (1997). Performance effects of motivational state: A meta-analysis. *Personality and Social Psychology Review*, 1, 170 – 182.

**Dr. Nantje Otterpohl**

Universität Bielefeld

Fakultät für Psychologie und  
Sportwissenschaft

Abteilung für Psychologie

Postfach 100131

33501 Bielefeld

[nantje.otterpohl@psychol.uni-giessen.de](mailto:nantje.otterpohl@psychol.uni-giessen.de)