

TBS-TK Rezension

Deutscher Rechtschreibtest für das erste und zweite/dritte und vierte Schuljahr, DERET 1-2+/3-4+

Prof. Dr. Barbara Gasteiger-Klicpera, Karl-Franzens-Universität Graz
Prof. Dr. Elisabeth Sticker, Universität Köln

Vorbemerkung

Die im folgenden ausgeführte Beurteilung bezieht sich vorwiegend auf die quantitative Auswertung des Tests. Eine Beurteilung der qualitativen Auswertung ist aufgrund fehlender Angaben zur Objektivität, Reliabilität und Validität der Analyse nicht leistbar. Auf dieses Problem wird an entsprechender Stelle hingewiesen.

Allgemeine Informationen über den Test, Beschreibung des Tests und seiner diagnostischen Zielsetzung

Die beiden DERET-Tests prüfen die Rechtschreibleistung von Schülerinnen und Schülern vom Ende der ersten bis zum Beginn der fünften Klasse. Sie liegen jeweils in zwei Parallelformen vor und sind als Einzel- oder Gruppentest gedacht. Die Kinder tragen diktierter Fließtexte und anschließend Lückenwörter in einen durch kleine Bildchen gut strukturierten Testbogen ein. Alle Wörter entstammen Rechtschreiblisten von Deutsch-Schulbüchern sowie Grundwortschätzen aus Lehrplänen. Die Tests dauern jeweils 30 bzw. 45 Minuten ohne bzw. mit Lückenwörtern. Für die Auswertung werden alle falsch geschriebenen Wörter aufsummiert. Neben der quantitativen kann eine

qualitative Auswertung mit der Erstellung eines individuellen oder klassenspezifischen Fehlerprofils vorgenommen werden. Durch Vergleich mit der Normstichprobe lassen sich für die Fehleranzahl Prozentränge sowie Prozentrangbänder ermitteln und Leistungsklassifizierungen vornehmen, die mithilfe von Ampelfarben veranschaulicht werden.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Die Autoren beziehen sich in Anlehnung an Treimann (1991, zitiert nach den Verfahrenshinweisen) auf das Modell der zweifachen Zugangswege beim Rechtschreiben, das einen direkten Zugang über das orthografische Lexikon von einem indirekten Zugang über die schrittweise Identifikation der Graphemfolge unterscheidet. Eine weitere wichtige Basis der Testkonstruktion stellt die Analyse der Lehrplaninhalte und der Schulbücher für das Fach Deutsch in der Grundschule dar. Kritisch anzumerken ist, dass die theoretischen Annahmen nicht mehr aufgegriffen werden, denn aus den Angaben zur Auswertung geht nicht hervor, wie die beiden unterschiedlichen Zugangswege identifiziert werden können und welche Implikatio-

nen dies für die Förderung haben könnte. Auch fehlen Hinweise darauf, welche Schwierigkeiten von Kindern (phonologische versus orthografische) zu welchen Arten von Fehlern führen. Unberücksichtigt bleibt auch die Art des Unterrichts, die den Erwerb bestimmter Rechtschreibstrategien ermöglicht und somit bestimmte Fehler nahelegt.

Objektivität

Der Test ist in seiner Durchführung weitestmöglich standardisiert. Die Durchführungsobjektivität ist durch die wörtliche Instruktion in der Durchführungsanleitung gegeben, sofern sich die Testleitung daran hält. Es finden sich konkrete Hinweise, wie bei Fragen zu reagieren ist und wie Abschreiben vermieden werden kann. Zu Beginn ist jeweils ein Übungssatz vorgesehen, damit die Kinder die Aufgabenstellung kennen lernen. Die Auswertungsobjektivität ist durch eindeutige Angaben, was als Fehler zu werten ist, gegeben. Eine Erleichterung bei der Auswertung stellt die farbige Darstellung der Prozentwerte dar, da damit eine rasche Einordnung der Ergebnisse möglich ist. Die Interpretationsobjektivität kann durch den Vergleich der Testergebnisse mit den auf der Normierungsstichprobe basierenden Normwerten ebenfalls als gegeben betrachtet werden. Hilfreich für die Interpretation sind Beispiele zu einer Einzel- und Klassenauswertung.

Normierung (Eichung)

Die DERET-Prozentrangnormen basieren auf 2095 Erst-, 2009 Zweit-, 2033 Dritt und 2021 Viertklässlern, untersucht im Jahre 2003. Die Normstichproben entstammen allen Bundesländern und entsprechen deren relativem Anteil von schulpflichtigen Kindern im Bundesgebiet. Im Jahre 2004 erfolgte eine Nacherhebung an 577 Erst-, 637 Zweit-, 501 Dritt- und 464 Viertklässlern mit einer erweiterten Testfassung zur Bestimmung der qualitativen Normwerte für verschiedene Fehlerarten. Allerdings liegt keine geschlechtsspezifische Normierung vor, was dadurch begründet wird, dass Unterschiede zwischen Jungen und Mädchen zwar vorhanden, aber nicht substanziell waren. Auch fehlen Angaben darüber, wie groß der Anteil an Kindern mit Migrationshintergrund in der Normierungsstichprobe ist; daher bleibt fraglich, ob der Test für Kinder mit Deutsch als Erst- und Zweitsprache gleichermaßen geeignet ist. Eine Konvertierungstabelle von der Prozentrangskala in andere Normskalen ist nicht enthalten.

Zuverlässigkeit (Reliabilität, Messgenauigkeit)

Es werden umfangreiche Daten zur Reliabilität der beiden Tests vorgelegt: Cronbachs α lag für alle vier Klassenstufen zwischen 0.89 und .93, der Halbierungskoeffizient nach Spearman-Brown zwischen .89 und .92.

Die Retest-Reliabilität (232 Erst-, 230 Zweit-, 190 Dritt- und 237 Viertklässler, Wiederholung nach 6 Wochen) betrug .81–.94, die Paralleltest-Reliabilität (73 Erst-, 79 Zweit-, 103 Dritt- und 74 Viertklässler mit unterschiedlicher Testform nach 4 Wochen) .87–.95.

Die Koeffizienten sind durchgängig mittelhoch bis hoch – wie in ähnlichen Testverfahren. Allerdings betreffen sie ausschließlich die quantitative, nicht aber die qualitative Auswertung. Insbesondere fehlen Angaben zur Interraterreliabilität. Es bleibt daher unklar, wie zuverlässig die qualitative Auswertung ist.

Gültigkeit (Validität)

Bei Rechtschreibtests, die das Schreiben von Wörtern aus dem Gedächtnis prüfen, liegt eine gewisse Augenscheinvalidität vor. Durch die Orientierung an den Lehrplänen aller Bundesländer bei der Testkonstruktion kann auch eine inhaltliche Validität angenommen werden. Weiterhin wurde die konvergente Validität überprüft. Die Korrelationen mit den Deutschen Rechtschreibtests 1-4 (allerdings schon 1990–1997 publiziert) liegen je Klassenstufe und Testform zwischen .63 und .83. Mit der Würzburger Leise-Leseprobe (WLLP, 1998, zitiert nach den Verfahrenshinweisen), bestehen Zusammenhänge von -.51 bis -.74 (negatives Vorzeichen, da umgekehrte Polung). Die Übereinstimmung mit dem Lehrerurteil (Kriteriumsvalidität) liegt bzgl. der Rechtschreibleistung zwischen $r = .58$ und .79, bzgl. der Leseleistung zwischen .59 und .73. Insgesamt sind die Validitätswerte (allerdings ausschließlich auf die quantitative Auswertung bezogen) als hoch einzustufen, d.h. die Rechtschreibleistungen können valide erfasst und interpretiert werden.

Weitere Gütekriterien (Störanfälligkeit, Unverfälschbarkeit und Skalierung)

Insgesamt ist der DERET ein aktuell normiertes und ökonomisches Verfahren. Das Problem des Abschreibens besteht trotz Vorsichtsmaßnahmen, ebenso könnte ein zu hoher Geräuschpegel während der Testvorgabe Kinder mit geringen Rechtschreibkompetenzen benachteiligen. Auf Verfälschungstendenzen in Richtung „faking bad“ (z. B. bei Hochbegabten) wird nicht eingegangen. Auch die Vorgabe durch unterschiedliche Testleiter kann zu Fehlern führen. Von Nachteil ist außerdem, dass zur Messgenauigkeit der qualitativen Analyse keine Daten vorliegen.

Die Einbeziehung von Fließtexten (zusätzlich zu Lückentexten) wird damit begründet, dass dadurch Kontexteffekte ausgeschaltet werden, weil die Kinder die Wörter dann nicht aufgrund des Kontextes identifizieren können. Dies ist eine neue Idee, die sicherlich für sich spricht.

Besonders ansprechend ist zudem die farbige Darstellung der Prozentwerte, womit eine rasche Einordnung der Probanden möglich ist. Dies hilft sicherlich auch, Auswertungsfehler zu vermeiden.

Abschlussbewertung/Empfehlung

Die DERET-Tests sind durch eine insgesamt hohe Testgüte gekennzeichnet. Gegenüber bereits vorliegenden Verfahren zur Prüfung der Rechtschreibleistung weisen sie eine Reihe an positiven Aspekten auf, z. B. Lehrplangültigkeit, Berücksichtigung eines Grundwortschatzes, aktuelle Normierung.

Anders als in bisherigen Rechtschreibtests werden nicht nur Lückenwörter, sondern auch Fließtexte diktiert. Dies wird durch eine höhere ökologische Validität begründet. Allerdings wirken die Fließtexte durchgängig infantilisierend und wenig zeitgemäß. Dies kann befremden und somit die Testergebnisse verfälschen.

Der Versuch eine qualitative Fehleranalyse zu ermöglichen und Normen bereitzustellen kann aufgrund fehlender Angaben zu den Gütekriterien nicht bewertet werden. Zudem – und das ist der kritischste Aspekt bei der Einschätzung dieses Tests – erfolgt diese Fehleranalyse nicht theoriegeleitet, sondern empiristisch. Daher bleibt unklar, wie die Fehler zustande gekommen sein könnten und welche Rechtschreibstrategien von den Kindern ver-

wendet werden. Folglich lassen sich Förderhinweise nur anhand der Fehlerschwerpunkte ableiten.

Zusammenfassend betrachtet handelt es sich um solide konstruierte Tests, die zur Erfassung der Rechtschreibfähigkeiten nicht nur im Querschnitt, sondern auch im Längsschnitt sehr zu empfehlen sind.

DERET 1-2+/3-4+	Die TBS-TK-Anforderungen sind erfüllt			
	voll	weitgehend	teilweise	nicht
Allgemeine Informationen, Beschreibung und diagnostische Zielsetzung	X			
Objektivität	X			
Zuverlässigkeit		X		
Validität	X			

Diese Testrezension wurde im Auftrag des Diagnostik- und Testkuratoriums der Föderation deutscher Psychologinnenvereinigungen (DGPs und BDP) gemäß den TBS-TK-Richtlinien (Testkuratorium, 2009, 2010) erstellt.

Testkuratorium (2009). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen. Revidierte Fassung vom 09. September 2009. *Report Psychologie*, 34, 470–478.

Testkuratorium (2010). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen. Revidierte Fassung vom 09. September 2009. *Psychologische Rundschau*, 61, 52–56.

Testinformationen

Stock, C. & Schneider, W. (2008). *DERET 1-2+*. Deutscher Rechtschreibtest für das erste und zweite Schuljahr. Göttingen: Hogrefe.

Stock, C. & Schneider, W. (2008). *DERET 3-4+*, Deutscher Rechtschreibtest für das dritte und vierte Schuljahr. Göttingen: Hogrefe.

Bezugsquelle: Testzentrale Göttingen, Herbert-Quandt-Str. 4, 37081 Göttingen.

DERET 1-2+, Test komplett bestehend aus: Manual, Durchführungsanleitung, 5 Testheften Klasse 1 Form A, 5 Testheften Klasse 1 Form B, 5 Testheften Klasse 2 Form A, 5 Testheften Klasse 2 Form B, 5 Auswertungsbogen Klasse 1 Form A, 5 Auswertungsbogen Klasse 1 Form B, 5 Auswertungsbogen Klasse 2 Form A, 5 Auswertungsbogen Klasse 2 Form B, 5 Klassenauswertungen und Mappe: 86,00 €.

DERET 3-4+, Test komplett bestehend aus: Manual, Durchführungsanleitung, 5 Testheften Klasse 3 Form A, 5 Testheften Klasse 3 Form B, 5 Testheften Klasse 4 Form A, 5 Testheften Klasse 4 Form B, 5 Auswertungsbogen Klasse 3 Form A, 5 Auswertungsbogen Klasse 3 Form B, 5 Auswertungsbogen Klasse 4 Form A, 5 Auswertungsbogen Klasse 4 Form B, 5 Klassenauswertungen und Mappe: 86,00 €.

Bitte zitieren Sie diesen Artikel wie folgt:

Gasteiger-Klicpera, B. & Sticker, E. (2011). TBS-TK Rezension: „Deutscher Rechtschreibtest für das erste und zweite/dritte und vierte Schuljahr, DERET 1-2+/3-4+.“ *Psychologische Rundschau*, 63, 75–77.

DOI: 10.1026/0033-3042/a000102