

Allgemeine vs. mathematische Begabung bei Fünftklässlern

Die MALU (Mathe-AG an der Leibniz Universität) fördert seit 2008 Fünftklässler, die einmal wöchentlich in Paaren selbständig Problemaufgaben bearbeiten. Neben Interesse ist eine nicht unterdurchschnittliche Begabung dafür eine plausible Voraussetzung. Dabei wird Begabung als eine stabile intrapersonale Eigenschaft aufgefasst, die interindividuelle Leistungsdifferenzen erklärt. Insoweit sie durch Tests erfassbar ist, kann sie als Prädiktor für zukünftige Leistungen und damit als ein Auswahlkriterium für die Förderung dienen. Über deren Verlauf und Ergebnisse wird an anderer Stelle berichtet – hier geht es um die Frage: Wie soll mathematische Begabung gemessen werden und welche Vorhersagen sind damit möglich?

Nach Heilmann (1999) sind prinzipiell drei verschiedene Modelle zur Erklärung mathematischer Leistung denkbar. Danach können interindividuelle Unterschiede mathematischer Leistung erklärt werden mit

1. unterschiedlich ausgeprägter mathematikspezifischer Begabung,
2. unterschiedlich ausgeprägter mathematikspezifischer und unterschiedlich ausgeprägter allgemeiner Begabung sowie
3. unterschiedlich ausgeprägter allgemeiner Begabung.

1. wird etwa von Krutetzki (1966) vertreten, 3. von Rost (2004, 43).

Um obige Frage für die MALU-Stichprobe zu beantworten, wurden an verschiedenen Hannoveraner Gymnasien ein mathematikspezifischer und ein allgemeiner Begabungstest mit 684 Fünftklässlern durchgeführt.

1. Allgemeine vs. mathematische Begabung

Allgemeine Begabung wird wie üblich verstanden als Intelligenz: „die Fähigkeit zum denkgestützten Lösen von Aufgaben und Problemen in Situationen, die für die Person neu und nicht allein durch Wissensabruf erfolgreich bearbeitbar sind, die Fähigkeit zum induktiv und deduktiv-logisch schlussfolgernden Denken, die Fähigkeit zum abstrakten Denken und die Fähigkeit zu Verständnis und Einsicht - zum Erkennen und zur Herstellung von Strukturen, Beziehungen, Sinnzusammenhängen und Bedeutungen.“ (Rindermann 2004, 374). Die Aufzählung suggeriert, dass Intelligenz verschiedenartige Aspekte hat, die durch die diversen Untertests von Intelligenztests operationalisiert werden – strittig war lange Zeit, ob diese manifesten Variablen besser als die unterschiedlichen Ausprägungen *einer* latenten Variable „allgemeine Intelligenz“ beschrieben werden können (Spearman's Generalfaktor *g*) oder durch *mehrere* latente Variable, die un-

terschiedlichen Fähigkeiten entsprechen (Thurstones „primary mental abilities“). Zur Klärung dieser Frage wurde die Faktorenanalyse entwickelt.

Bedenken, ob Intelligenztestaufgaben solche Fähigkeiten messen, die wesentlich für mathematische Begabung sprechen, werden vielfach von Mathematikdidaktikern geäußert. Geschlossene, zielgerichtete Intelligenztestaufgaben ließen lediglich eine einzige richtige Lösung zu, klammerten dabei Originalität und Einfallsreichtum aus und transportierten die Vorstellung von Mathematik als wohldefiniertem Gegenstand (Käpnick 1998, 68).

Ausgehend von dem Vorverständnis von *mathematischer Begabung* eigener Beobachtungen von mathematisch begabten Kindern und den Merkmalsystemen insbesondere von Krutetzki (1966) und Kießwetter (1985) stellt Käpnick in einem eigenen System „Merkmale für die Erfassung von Dritt- und Viertklässlern mit einer potentiellen mathematischen Begabung“ zusammen (Käpnick 1998, 119). Dazu zählt er sowohl *mathematikspezifische Begabungsmerkmale* wie Originalität und Phantasie bei mathematischen Aktivitäten, Gedächtnisfähigkeit, Fähigkeit zum Strukturieren, Fähigkeit zum Wechseln der Repräsentationsebenen, Fähigkeit zur Reversibilität und zum Transfer als auch *begabungsstützende allgemeine Persönlichkeitseigenschaften* wie hohe geistige Aktivität, Anstrengungsbereitschaft, Freude am Problemlösen und Beharrlichkeit.

2. Auswahl der Tests

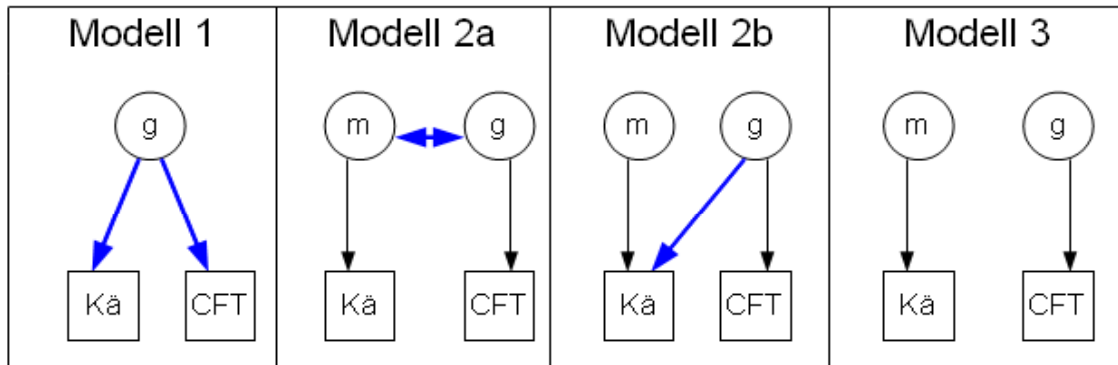
Als mathematischer Begabungstest wurde der von Käpnick zusammengestellte *Indikatoraufgaben-Test* (Käpnick 1998) in einer verkürzten Version¹ eingesetzt. Als IQ-Test wurde der CFT-20R ausgewählt. Wie Ravens APM ist der CFT-20R unabhängig vom Sprachverständnis und erfasst die „Fähigkeit zur Problemerkennung in neuartigen Situationen“ (Weiß 2006, 16). Im Gegensatz zum APM existieren für den CFT-20R aktuelle deutsche Normwerte für Fünftklässler. Zudem konnten von uns keine Deckeneffekte beobachtet werden. Andere IQ-Gruppentests mit Normdaten für Fünftklässler schieden aufgrund ihrer Zeitdauer von vornherein aus.

3. Ergebnisse

Von 684 Fünftklässlern liegt sowohl das CFT-20R- als auch das Käpnick-Testergebnis vor. Beide Punktzahlen korrelieren zu $r=0.374^{**}$ miteinander, so dass von einem schwach positiven linearen Zusammenhang ausgegangen werden kann. Anders formuliert: 14% der Varianz des einen Testergebnisses kann durch die des anderen erklärt werden. Da ein direkter kausaler Zusammenhang der Testpunktzahlen theoretisch ausgeschlossen wer-

1 Details und Literatur in Langfassung auf www.idmp.uni-hannover.de unter Gawlick bzw. Lange

den kann, sind entsprechend den Heilmannschen Fällen folgende Deutungsmodelle zur Erklärung der Korrelation der manifesten Testvariablen durch unterschiedliche Wirkzusammenhänge der latenten Begabungsvariablen (allgemeine: g, mathematische: m) denkbar.



Die dazugehörigen Strukturgleichungsmodelle wurden einer konfirmatorischen Faktorenanalyse mit AMOS 17 unterzogen. Alle Indizes¹ deuten darauf hin, dass Modell 2b die empirischen Relationen zwischen den manifesten Variablen am besten erklären kann (vgl. Brunner 2006, 100). Dieses Ergebnis stimmt gut damit überein, dass die Felder 12 und 21 in nachstehender Vierfeldertafel² besetzt sind – wenn auch signifikant geringer als für unabhängige Merkmale zu erwarten ($\chi^2=30,618$, 2-seitig, $p<0,01$).

		Käpnick-Test	
		1	2
IQ-Wert	1	Anzahl	580
		Erwartete Anzahl	568,3
	2	Anzahl	41
		Erwartete Anzahl	52,7

Um Aussagen über die prädiktive Validität der beiden Tests machen zu können, wurden von einer Sechstklässlerstichprobe die Mathematik- und die Deutschhalbjahresschulnoten sowie die Einzelbearbeitungen von 4 ausgewählten MALU-Aufgaben erhoben und bivariat bzw. multipel korreliert:

		Deutschnote	Mathenote	Indiv. Bearbeitungen
		N = 130 Kl. 6	Käpnicktest	$r=0.350^{**}$
CFT	$r=0.199^{**}$		$r=0.394^{**}$	$r=0.305^{**}$
Käpnicktest+CFT	--		$R=0.469$	$R=0.347$
Käpnicktest+CFT+Mathenote	--		--	$R=0.399$

2 2 = weit überdurchschnittlich gute Testleistung; 1 = Rest. Als weit überdurchschnittlich intelligent gelten nach Rost(2004) Kinder, die mindestens einen IQ von 130 erreichten (d.h. mindestens zwei Standardabweichungen über dem Mittelwert der Normstichprobe). Der gleiche Anteil der Probanden soll auch als mathematisch hochbegabt eingestuft werden. Die tatsächliche Probandenzahl weicht davon leicht ab, da die Cut-Off-Punktzahl im Käpnick-Test von mehreren Fünftklässlern erreicht wurde.

4. Diskussion der Ergebnisse

Sowohl der Intelligenz- als auch der Käpnick-Test sind valide Prädiktoren, die allerdings jeweils nur ca. 16% Varianzaufklärung für mathematische Schulleistungen und ca. 9% für mathematisches Problemlösen erbringen. Es ist überraschend, dass letzterer Wert deutlich geringer ausfällt.

Die Korrelationen der Testergebnisse mit den Noten liegen im Rahmen des üblichen: Der CFT-20R korreliert laut Handbuch in den Gymnasialklassen 5-9 zu .40 mit der Mathematik-Note (Weiß 2006, 87). Für den Käpnick-Test sind uns keine Vergleichsdaten bekannt.

Die Interkorrelation der Tests liegt niedriger als z.B. die Interkorrelation der Ergebnisse des TIMSS-Mathematiktests mit kognitiven Grundfähigkeiten (figural: .49, verbal: .59, vgl. Baumert et al. 1997). Dies deutet darauf hin, dass das hier gemessene Konstrukt weiter von allgemeiner Begabung entfernt ist und damit mathematikspezifischer sein könnte. Allerdings liegt die Varianzaufklärung für Problemaufgaben noch unter der für Schulnoten.

Durch Kombination der Tests wird aufgrund der Interkorrelation die prädiktive Validität zwar deutlich erhöht, die Varianzaufklärung bleibt aber unbefriedigend. Die multiple Korrelation der Mathematikhalbjahresnote mit Käpnick- und CFT-Test beträgt $R=0.469$, also $R^2=0.220$. Für die Problemaufgaben liegen die Werte mit $R=0.347$ und $R^2=0.120$ noch niedriger, so dass die Vermutung nahe liegt, dass hierbei ein wesentlicher Faktor nicht miterhoben wurde – welcher das sein könnte, bleibt jedoch unklar. Es ist nicht das mathematische Vorwissen, denn die multiple Korrelation steigt lediglich auf $R=0.399$, wenn die Mathematikhalbjahresnote als zusätzlicher Prädiktor aufgenommen wird. Auch die Sprachkompetenz kann ausgeschlossen werden, da sich die multiple Korrelation bei Hinzunahme der Deutschnote lediglich um 0.014 auf $R=0.413$ erhöht. Insgesamt erklären die beiden Tests und die beiden Noten zusammen lediglich 17% der Varianz der individuellen Aufgabebearbeitungen.

Dass in anderen Förderprojekten die Ergebnisse der Förderung ebenfalls nur unzureichend durch Tests vorhersehbar sind, stützt die in der Community verbreitete Präferenz für eine ganzheitliche Prozessdiagnostik, die allerdings i.d.R. auch Tests mit einbezieht.

Wenn aus ökonomischen Gründen nur einer der beiden Tests als ein Auswahlkriterium fungieren soll, könnte man den Käpnick-Test bevorzugen, da dieser auch auf den g-Faktor lädt. Wenn man auch hohe Leistungen, die nur aufgrund hoher allgemeiner Begabung erbracht werden, miterfassen möchte, braucht man beide Tests.

Literatur¹