

Mal rauf, mal runter

Wie stabil ist Intelligenz?

Die Befunde wurden in der Presse heiß diskutiert: Obwohl es immer wieder heißt, dass das kognitive Fähigkeitsniveau etwa ab dem Grundschulalter stabil ist, zeigte eine kürzlich veröffentlichte Studie, dass die Intelligenz auch im Jugendalter teilweise massiv schwanken kann. Ob dieser Brei wirklich so heiß gegessen werden muss, wie er gekocht wurde, fragt sich Tanja in der heutigen Jubiläumsausgabe der „Streifzüge“.

Wann sollte man ein Kind auf Hochbegabung testen lassen? Stabil ist die Intelligenz etwa ab dem Grundschulalter – so der weitgehende Konsens unter Diagnostikern. Deutlich jünger als fünf sollten die Kleinen nicht sein, auch wenn es Tests gibt, die sogar ab zweieinhalb Jahren angewandt werden können; denn je jünger ein Kind ist, desto fehleranfälliger ist ein solcher Test. Entsprechend weisen die Testhandbücher darauf hin, wie wichtig eine gute Beziehung in der Testsituation ist, um das Beste aus dem Kind herauszuholen und ein aussagekräftiges Ergebnis zu erhalten, das tatsächlich kognitive Fähigkeiten erfasst – und nicht die Fähigkeit, mit einschüchternden Situationen umzugehen.

Ungenauere IQ-Tests?

Kein Intelligenztest misst hundertprozentig genau – Angst ist nur einer von vielen Gründen, warum der „wahre“ IQ

vielleicht doch etwas höher liegt (wobei auch diejenigen, mit denen der Test normiert wurde, vermutlich nicht völlig entspannt waren ...). Rateglück kann das Ergebnis etwas steigern; wer es also auf ein möglichst hohes (und nicht auf ein möglichst akkurates) Resultat anlegt, sollte im Zweifel lieber irgendeine als gar keine Antwort ankreuzen. Diesen leichten Ungenauigkeiten begegnet man, indem man ergänzend zum gemessenen Punktwert ein Intervall angibt, in dem der „wahre“ IQ mit einer bestimmten Wahrscheinlichkeit liegt. Bei 95-prozentiger Sicherheit liegt dieses beim HAWIK-IV (seit 2011 WISC-IV), dem bei Kindern gängigsten Test, für den Gesamt-IQ bei etwa plus/minus fünf Punkten, für die Teilbereiche etwas höher.

Ein Gütekriterium von IQ-Tests ist, dass sie kognitive Fähigkeit stabil erfassen; bei wiederholter Messung sollten die Ergebnisse also ähnlich ausfallen. Diese Stabilität (oder „Retest-Reliabilität“) wird in der Regel in Form einer Korrelation* berichtet. Wer beim ersten Mal gut abschneidet, sollte auch beim zweiten Mal oben mit dabei sein. Über das absolute IQ-Niveau (also das Verhältnis zur Norm) sagt sie nichts aus, sondern nur über das relative Niveau im Vergleich zum Rest der Stichprobe. Eine perfekte Korrelation läge also auch vor, wenn sich alle zum zweiten Messzeitpunkt um 20 IQ-Punkte verbessert hätten. Je kleiner die Stichprobe, desto stärker machen sich „Ausreißer“ bemerkbar, die aus dem Muster herausfallen; größere Stichproben sind weniger verzerrungsanfällig.

Neue Studie – neue Erkenntnisse?

Was hat es nun mit der eingangs erwähnten Studie auf sich? Kurz zusammengefasst: Die Autoren testeten den

* Eine Korrelation ist ein standardisiertes Maß dafür, wie eng zwei Merkmale (hier: Intelligenz zum ersten und zweiten Messzeitpunkt) statistisch zusammenhängen. Sie reicht von -1 (perfekt negativer Zusammenhang) bis +1 (perfekt positiver Zusammenhang); eine Korrelation von 0 bedeutet, dass kein systematischer Zusammenhang vorliegt.

Streifzüge durch die Begabungsforschung (XXV)

IQ von Jugendlichen im Alter zwischen 12 und 16 Jahren und untersuchten außerdem die Struktur des Gehirns mit Hilfe von Magnetresonanztomografie. Diese Untersuchung wurde nach drei bis vier Jahren noch einmal wiederholt; hier waren die Jugendlichen zwischen 15 und 20 Jahren alt. Beim ersten Mal wurde das englischsprachige Äquivalent zum HAWIK-III (damals der aktuelle Intelligenztest für Kinder und Jugendliche), beim zweiten Mal der WISC-III (die Erwachsenenversion) angewandt, die sich in der damaligen Fassung in einen Verbal- und einen Handlungs-IQ sowie einen Gesamt-IQ gliederten**. Im direkten Vergleich zeigten sich teilweise starke Veränderungen in den drei Bereichen – bis hin zu 23 IQ-Punkten! Diese waren auch in unterschiedlichen Arealen des Gehirns in Form von höherer Dichte der grauen Substanz*** nachweisbar: Veränderungen im Verbal-IQ zeigten sich vor allem in dem Teil, der durch die Artikulation von Sprache aktiviert wird, Veränderungen im Handlungs-IQ hingegen in dem Bereich, der für Bewegungen der Hand zuständig ist. In den Arealen, die früheren Studien zufolge für die allgemeine Intelligenz zuständig sind, fanden sich hingegen keine substantiellen Veränderungen.

Kritik an der Repräsentativität

Der absolute IQ kann sich im Jugendalter sowohl insgesamt als auch in seinen Teilbereichen verändern, letzteres korreliert mit Veränderungen in unterschiedlichen (wenngleich sehr kleinen) Hirnarealen – soweit die Befunde der Untersuchung. Schaut man sich die Studie aber genauer an, bekommt man Zweifel, ob die Resultate tatsächlich so generalisierbar sind, wie das die Presse

gerne hätte: Mit 33 von ursprünglich 47 Probanden, für die Daten aus beiden Messzeitpunkten vorliegen, ist die Stichprobe klein. Und sehr heterogen – allein die IQ-Spannbreite betrug bei der ersten Messung zwischen 77 und 135, bei der zweiten Messung zwischen 87 und 143; insgesamt lag der Mittelwert im oberen Durchschnitt (IQ 112 beziehungsweise 113). 18 der Probanden waren mit einer Lese-Rechtschreib-Schwäche (LRS) und teilweise noch anderen Schwierigkeiten diagnostiziert, 11 waren auf Grundlage eines Einstufungstests für eine Spezialschule als Hochleister identifiziert worden. Ein breites Fähigkeitsspektrum in Ehren – aber sind die ganz normalen Schüler nicht ein wenig unterrepräsentiert?

Eine Frage der Interpretation

Kritisch ist auch die Interpretation der IQ-Ergebnisse. Zwar sind die beiden Wechsler-Tests ähnlich strukturiert und gegliedert; ob sie tatsächlich äquivalent sind oder ob verschiedene Subtests in verschiedenen Entwicklungsphasen nicht doch unterschiedliche Aspekte der Intelligenz messen, wird hier (wie in den meisten Studien dieser Art) nicht

** Diese Unterteilung ist nicht theoretisch begründet; Wechsler selbst entschied sich aus rein praktischen Gründen für die Zweiteilung in Verbal- und Handlungsteil, räumte aber ein, dass dies keineswegs die einzige Möglichkeit sei, Intelligenz zu strukturieren. Die neueste Fassung, der WISC-IV, gliedert sich in vier statt bisher zwei Teilbereiche.

*** Graue Substanz besteht größtenteils aus Nervenzellkörpern, weiße Substanz hingegen aus den neuronalen Leitungsbahnen (Axonen).

Über die Autorin

Dr. Tanja Gabriele Baudson arbeitet als Hochbegabungs- und Kreativitätsforscherin an der Universität Trier. Sie ist Beisitzerin für Hochbegabtenforschung von Mensa in Deutschland e. V. Alle bisherigen Artikel dieser Reihe stehen auf ihrer Website bei der Uni Trier zur Verfügung. Erreichen könnt ihr Tanja per Mail.

Link und Mail

- ▶ <http://www.uni-trier.de/index.php?id=20275>
- ▶ forschung@mensa.de

„Letztendlich geht es bei Bildung doch immer darum, einen Menschen dort abzuholen, wo er zum jeweiligen Zeitpunkt steht.“

gezeigt. Die relativ zum Verbal-IQ deutlich niedrigeren Korrelationen zwischen dem Handlungsteil des Kinder- und des Erwachsenentests könnten darauf hinweisen.

Im Mittel sind die Veränderungen auch nicht so groß, wie die Extreme vermuten lassen: Für den Gesamt-IQ und die Teilbereiche liegen sie zwischen -1 und +3 Punkten bei hoher Retest-Reliabilität – ein Großteil dürfte innerhalb des Vertrauensintervalls liegen. Wollten die Autoren den in der Presse verbreiteten Punkt fokussieren, dass Intelligenz auch im Jugendalter noch veränderbar ist, wären differenzielle Analysen der aussagekräftigere Weg – zum Beispiel Veränderungen in Abhängigkeit von Alter, einer vorhandenen LRS, vom IQ (Stichwort Regression zur Mitte****). Ganz zu schweigen von Bildungs- und Fördereinflüssen, denn was ist in der Zwischenzeit mit den Jugendlichen überhaupt passiert? Die Korrelationen zu Veränderungen im Gehirn sind interessant und neu; jedoch handelt es sich um sehr kleine und spezifische Areale, während die Bereiche, die üblicherweise mit allgemei-

ner Intelligenz in Verbindung gebracht werden, keine signifikante Veränderung zeigten. Die Überprüfung erfolgte übrigens nicht hypothesengeleitet für bekannte Areale, sondern explorativ für das gesamte Gehirn. Vielleicht war die Kinderversion des Tests für Probanden an der oberen Altersgrenze auch schon zu leicht (aufgrund solcher „Deckeneffekte“ sollte er etwa ab zwölf Jahren auch nicht mehr zur Hochbegabungsdiagnostik verwendet werden). Ein „Sinken“ der Intelligenz bei Verwendung des Erwachsenentests wäre in dem Fall nicht ungewöhnlich; dies wird jedoch nicht aufgeschlüsselt. Und nicht zuletzt drängte sich mir auch die Frage nach dem Motivationsniveau der zum Teil ja noch mitten in der Pubertät befindlichen Jugendlichen auf ...

Ernüchterndes Fazit

Die Ergebnisse sind weit weniger sensationell als in der Presse dargestellt, und die Folgerung, nicht zu früh zu selektieren, ist in Anbetracht der Notwendigkeit individueller Förderung ein Gemeinplatz (den man mit Blick auf die politische Realität aber auch gerne wiederholen darf). Letztendlich geht es bei Bildung doch immer darum, einen Menschen dort abzuholen, wo er zum jeweiligen Zeitpunkt steht.

Tanja Gabriele Baudson

Literatur

- ▶ Ramsden, S. et al. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature* (online publiziert am 19.10.2011). doi: 10.1038/nature10514
- ▶ Wechsler, D. (2011). *WISC-IV*. (Deutsche Bearbeitung von F. und U. Petermann.) Frankfurt/Main: Pearson.

**** Die Regression zur Mitte besagt hier, dass Personen, die bei der ersten Testung sehr hohe oder niedrige Werte erzielt haben, aufgrund des Messfehlers bei der zweiten Messung eher Werte in Richtung des Mittelwerts erzielen.