

Intelligenztest

aus Wikipedia, der freien Enzyklopädie

Ein **Intelligenztest** ist ein Instrument der psychologischen Diagnostik zur Messung der Intelligenz einer Person. Da Intelligenz und ihre Bereiche unterschiedlich definiert werden können, gibt es sehr verschiedenartige Intelligenztests. Ziel der Messung von Intelligenz ist die Beschreibung von Personen hinsichtlich ihrer kognitiven Fähigkeiten. Dabei geht man davon aus, dass Leistungsunterschiede in Intelligenztests auch Unterschiede der Leistungsfähigkeit im modernen täglichen Leben abbilden. Psychologische Tests werden als Hilfsmittel zum Finden einer optimalen Entscheidung im Rahmen verschiedener diagnostischer Fragestellungen eingesetzt. Die Erfassung der Intelligenz kann z. B. für die Vorhersage von Berufserfolg oder Berufseignung (Personalauswahl oder Berufsberatung) oder für Empfehlungen einer Schulausbildung oder einer Studienwahl notwendig sein. Auch verschiedene klinische Fragestellungen (Vorliegen von Erkrankungen wie Demenz oder Intelligenzminderungen oder die Anwendung beruflicher Rehabilitationsmassnahmen) können die Erfassung der Intelligenz notwendig machen.

Am bekanntesten ist immer noch das Ergebnis einiger dieser Tests, der Intelligenzquotient (IQ). Umgangssprachlich werden Intelligenztests manchmal noch **IQ-Tests** genannt. Wegen der Gefahr der Verabsolutierung des IQ als *Etikett* einer Person (wie Körpergröße oder Gewicht - also unabhängig von einer konkreten Fragestellung) verwendet man diese Größe in der Fachsprache nicht mehr und greift auch auf andere Normskalen zur Beschreibung von Intelligenz-Leistungen zurück (z. B. Standardwerte SW).

Normskalen für die Intelligenzmessung

Man geht davon aus, dass Intelligenz normal verteilt ist. Der Intelligenzquotient (IQ) wird als Abweichungsquotient der Rohwerte zum Mittelwert in Einheiten der Standardabweichung von 15 mit einem Mittelwert von 100 angegeben.

$$IQ = 100 + 15 \cdot \frac{(x - \mu)}{\sigma} \quad (\text{HAWIE und HAWIK Hamburg-Wechsler-Intelligenztest})$$

Eine andere gebräuchliche Größe sind die Standardwerte (SW), die manchmal mit dem IQ verwechselt werden. Hier ist der Mittelwert ebenfalls 100, die Standardabweichung aber 10.

$$SW = 100 + 10 \cdot \frac{(x - \mu)}{\sigma} \quad (\text{IST Intelligenz-Struktur-Test nach Rudolf Amthauer})$$

Beim Vergleich der Werte verschiedener Intelligenztests ist zu beachten, dass unterschiedliche Leistungsanforderungen (Validitätsbereiche) die Aussagefähigkeit der Tests ebenso bestimmen wie die Verwendung unterschiedlicher Normierungsstichproben (z. B. unterschiedliche Repräsentativität für die Gesamtbevölkerung), auf die sich der gemessene Wert für Intelligenz bezieht.

Arten von Intelligenztests

Intelligenztests sind jeweils auf eine Intelligenztheorie bezogen. Entsprechend gibt es Tests zur Erfassung einer (einzigen) allgemeinen Intelligenz oder von Komponenten (Faktoren, Dimensionen) der Intelligenz.

Allgemeine Intelligenz

Bezieht sich auf das Generalfaktormodell von Charles Spearman und wird beispielsweise durch den Zahlenverbindungstest (ZVT) von Erwin Roth oder die Matrizentests von John C. Raven erfasst.

Kristalline (kristallisierte) und fluide Intelligenz

Diese Gegenüberstellung geht auf den britisch-amerikanischen Psychologen R.B.Cattell zurück und bedeutet die Unterscheidung von Problemlösefähigkeit (Fähigkeit, Problemstellungen zu lösen und sich an neue Problemstellungen anzupassen) - als *fluide Intelligenz* und Wissen (die früher erworbenen Wissensbestände als bildungsabhängige Komponente) - als *kristalline Intelligenz*. Weist eine Person unterschiedliche Ausprägungen in beiden Bereichen auf, können sowohl in der Klinischen Psychologie (z. B. Abbaudiagnostik) als auch in der Entwicklungspsychologie (z. B. Liegt ein Fähigkeit entsprechender Bildungsstand vor?) wichtige Schlüsse gezogen werden.

Verbale und praktische (Handlungs-) Intelligenz

Diese Unterteilung beruht auf dem Intelligenzkonzept von David Wechsler, der Intelligenz als „die zusammengesetzte oder globale Fähigkeit des Individuums (bezeichnet,) zielgerichtet zu handeln, rational zu denken und sich wirkungsvoll mit seiner Umwelt auseinanderzusetzen“ (Wechsler, 1956, S. 13.)^[1]. Wechsler legt dabei fest, dass Intelligenz aus Verbal- und Handlungsintelligenz bestehe, was seinen Niederschlag in den auf seinem Konzept beruhenden Wechsler-Intelligenzskalen gefunden hat. Anders als in den aktuellen Versionen der Tests (WIE – Wechsler Intelligenztest für Erwachsene; HAWIK©-IV – Hamburg-Wechsler-Intelligenztest für Kinder, 4. Auflage) bestand bislang eine strikte Trennung der Untertestaufgaben in einen so genannten Verbal- vs. Handlungsteil. In den beiden neuesten Auflagen hingegen wurde diese Zweiteilung zugunsten einer Differenzierung in vier so genannte Indices aufgegeben, die nunmehr die Leistungsfähigkeit in den vier Bereichen Sprachverständnis, Wahrnehmungsgebundenes logisches Denken, Arbeitsgedächtnis und Verarbeitungsgeschwindigkeit beschreiben.

Im Rahmen des Tests für Kinder gelten folgende Zuordnungen:

- Sprachverständnis
 - Gemeinsamkeiten finden
 - Wortschatztest
 - Allgemeines Verständnis
 - optional Allgemeines Wissen

- Wahrnehmungsgebundenes logisches Denken
 - Mosaiktest
 - Bildkonzepte
 - Matrizentest
 - optional Bilderergänzen

- Arbeitsgedächtnis
 - Zahlennachsprechen
 - Buchstaben-Zahlen-Folgen
 - optional Rechnerisches Denken
- Verarbeitungsgeschwindigkeit
 - Zahlen-Symbol-Test
 - Symbolsuche
 - optional Durchstreichtest

Mehrfaktorenkonzepte

beruhen z. B. auf Modellen von Louis Leon Thurstone oder Adolf Otto Jäger. *Beispiel:*

- I-S-T 2000R (dieser Test verwendet anstatt des IQ den Standardwert SW)
 - verbale Kompetenz (Satzergänzung, Analogien, Gemeinsamkeiten)
 - numerische Kompetenz (Rechenaufgaben, Zahlenreihen, Rechenzeichen)
 - figurale Kompetenz (Figurenauswahl, Würfelaufgaben, Matrizen)
 - Merkfähigkeit, Schlussfolgerndes Denken (Reasoning)
 - *Erweiterungsmodul:* fluide und kristallisierte Intelligenz (Generalfaktoren)

Intelligenztests und Bildungsevaluation

Nach den Ergebnissen von Rindermann (2006^[2]) sollen auch die bei den PISA-Studien, bei TIMSS und bei IGLU eingesetzten Testverfahren als Intelligenztests zu verstehen sein – dies ist allerdings umstritten (Baumert u. a. 2007^[3]). Die Validität und Verlässlichkeit der PISA-Tests als Maß des Allgemeinen Faktors der Intelligenz wird von der Studie als gleichwertig oder sogar höher als bei den oben angeführten Tests angesehen. Allerdings werden PISA-Tests nicht unter ausreichend abgesicherten Bedingungen durchgeführt, so dass die Ergebnisse in der Regel nicht für eine Intelligenzaussage zu den einzelnen Schülern verwendet werden können.

Nichtsprachliche Intelligenztests

Es gibt eine Reihe von Voraussetzungen, deren Nichtbeachtung zu intuitiv falschen Ergebnissen bei IQ-Tests führt. Sprachlastige IQ-Tests etwa setzen voraus, dass die Person die Sprache kennt, in der die Fragen vorgelegt werden: Kinder bestimmter Gruppen von Gastarbeitern und Immigranten schneiden dabei oft schlecht ab ^[4].

Für diese Fälle gibt es sprachfreie IQ-Tests wie zum Beispiel SPM (Raven Standard Progressive Matrices)^[5] und den CFT 20 (*Culture Fair Test*)^[6]. Der CFT 20 kann auch von Menschen mit geringen kulturtechnischen Kompetenzen angewendet werden ^[7]. Der SON-R 5 1/2-17 und der Bochumer Matrizentest (BOMAT) sind ebenfalls Beispiele für sprachfreie Intelligenztests.

Konstruktion

Die Autoren von Intelligenztests konstruieren ihre Tests nach impliziten Annahmen bzw. in Anlehnung an bestehende Intelligenztheorien. Es kann davon ausgegangen werden, dass verschiedene Intelligenztests unterschiedliche Ausschnitte der Intelligenz erfassen.

Einige Verfahren verwenden sehr verschiedene Untertests/verschiedene Itemformen (IST2000, BIS), während andere sehr homogene aber besonders repräsentative Itemformen verwenden (*Mental Speed*, Matrizen tests).

Einige Verfahren sind besonders für die Erfassung der allgemeinen Intelligenz konzipiert (Modell des Generalfaktors der Intelligenz) während andere (WIT, LPS) eher an der spezifischen Ausprägung der einzelnen Faktoren der Intelligenz interessiert sind.

Gütekriterien der Tests

Gängige Intelligenztests erreichen meist eine gute Reliabilität. Ebenso sind die Kriteriumsvalidität und die prädiktive Validität bezogen auf Schulnoten meist gut. Dies liegt daran, dass die Gültigkeit der Tests von vornherein danach konzipiert wird, welche Unterschiede zwischen Schülern höher- oder geringwertiger Allgemeinbildung an verschiedenen Schultypen aufgefunden werden können. Langfristig bewährt hat sich ein Test demnach vor allem dann, wenn der Schulerfolg mit dem Abschneiden in IQ-Tests korreliert. Die darin liegende Unterstellung, es handele sich bei der Intelligenz um ein getrennt von Testkonstruktionen existierende und aufzufindende allgemeine Fähigkeit, die sich so gesehen letztlich am Schulerfolg zeigt, wird von den Testkonstruktoren selbst problematisiert. Die Lösung der Frage, ob also Intelligenztests messen was sie messen sollen und wie gut, beantwortet damit zugleich die insoweit dann nicht weiter zur Debatte stehende Frage nach dem begrifflichen Inhalt von Intelligenz - eben im Sinne der Anerkennung von Intelligenztests.^[8]

Die Vorhersagekraft von Intelligenztests ist oft nicht völlig unabhängig von der Höhe des Testwertes. Bei sehr niedrigen oder sehr hohen IQ-Werten ist die Vorhersagekraft (z. B. für Berufserfolg) meist etwas größer als im mittleren Bereich.

Der Einsatz von Intelligenztests bei der Bewerberauswahl, z. B. im Rahmen eines Assessment-Centers, hat immer unter Aufsicht eines Diplom-Psychologen zu erfolgen, anderenfalls kann das Verfahren angefochten werden.

Normierung

Intelligenztests werden nach Altersgruppen und gegebenenfalls nach anderen Merkmalen, wie Schulabschlüssen normiert. Manche Autoren sprechen von der so genannten Eichstichprobe. Diese Normierung erfolgt getrennt in den einzelnen Ländern. Länderübergreifende Vergleiche sind daher nur eingeschränkt möglich.

Der Flynn-Effekt bezeichnet die Tatsache, dass bis in die 1990er Jahre die Ergebnisse von IQ-Tests im Mittel höhere Werte erbrachten, die gemessene Intelligenz also offenbar zunahm. Mit Beginn der 1990er-Jahre stagnierte der IQ, und seit dem Ende der 1990er Jahre nimmt er sogar wieder ab^{[9][10][11]}. Dadurch ist es notwendig, Intelligenztests regelmäßig neu zu eichen. Bis 1990 mussten sie komplizierter gemacht werden, jetzt müssen sie wieder einfacher gemacht werden.

Kritik

Kinder, die bei Intelligenztests gut abschneiden finden sich vor allem in den oberen Schichten ^[12]. Es wird diskutiert, ob dies daran liegt, dass traditionelle Intelligenztests gegenüber Kindern anderer Schichten unfair sind. Anastasi berichtet etwa, dass bei den traditionellen Intelligenztests Ober- und Mittelschichtskinder stets besser abschneiden als Arbeiter- und Unterschichtskinder. Jedoch gibt es auch Testverfahren, bei denen Unterschichtskinder anderen voraus sind. Dazu gehört etwa der Minnesota-Test für mechanisches Zusammenbauen (*Minnesota Mechanical Assembly Test*). Dieser Test ist besonders gut geeignet, um mechanische Fähigkeiten zu testen. Den Probanden wird ein T-Score zugeordnet. Je höher dieser ist, desto besser sind die Fähigkeiten. Die Mittelwerte dieses Tests weisen keine konsistente Beziehung zum Sozialstatus auf. Die besten Werte erreichen jedoch Unterschichtskinder. Als eine mögliche Ursache wird angesehen, dass Unterschichtssöhne mehr Erfahrung im Umgang mit mechanischen Objekten hinter sich haben und somit mit den Aufgaben dieses Tests vertrauter sind ^[13].

Wegweisend ist hier die Studie von Janke und Havighurst, die bereits 1945 stattfand. Diese testeten Kinder und Jugendliche aus „einer typischen Stadt im mittleren Westen der USA“ mit verschiedenen Intelligenztests ^[14]. Bei den meisten Tests schnitten Unterschichtskinder deutlich schlechter ab als Kinder der Ober- oder Mittelschicht. Nicht jedoch beim Minnesota Mechanical Assembly Test ^[15].

Dies wird hier nicht für alle Tests in aller Ausführlichkeit dargestellt werden. Stattdessen werden hier nur die Ergebnisse zweier Tests dargestellt werden: Die des Stanford-Binet-Test (ein Test, bei dem es sehr starke Schichtunterschiede gibt) und die des Minnesota Mechanical Assembly Test. Folgendes Ergebnis zeigte sich bei 16-jährigen Schülern aus verschiedenen Herkunftsschichten ^[16]:

Sozialschicht der Eltern	Ergebnis auf dem Stanford-Binet (IQ)	Ergebnis auf dem Minnesota Mechanical Assembly Test (T-Score)
Alle Sozialschichten zusammen	108 dies deutet darauf hin, dass Leute mit sehr niedrigem IQ schon nicht mehr in der Stichprobe waren, siehe dazu unten	50
A&B (die höchsten beiden Herkunftsschichten)	128	46,8
C	112	51,6
D	104	48,8
E (niedrigste Herkunftsschicht)	98	53,0

Hier zeigte sich also, dass die Jugendlichen aus der höchsten Sozialschicht den höchsten IQ (gemessen durch den Stanford-Binet) hatten. Mit sinkender Sozialschicht nahm der IQ ab. In den beiden untersten Sozialschichten (D&E) waren zum Zeitpunkt der Testung einige Jugendliche bereits nicht mehr in der Schule^[17]. Es wird vermutet, dass es sich bei diesen um Jugendliche mit besonders niedrigem IQ (gemessen auf dem Stanford-Binet) handelt, weil Highschool-Drop-Outs meistens einen niedrigen IQ haben. Die IQs in Gruppe D&E sind wahrscheinlich noch niedriger^[18]. Der T-Score hingegen folgte nicht diesem Muster. Er war in der untersten Schicht am höchsten. Da der T-Score aller Sozialschichten beim Minnesota Mechanical Assembly Test 50 war (und dies dem erwarteten Wert entspricht), ist bei diesem Test nicht mit systematischen Verzerrungen zu rechnen.

Erstmals konnte in einer Studie ein Verfahren demonstriert werden, mit dem die Probanden deutlich besser im Test abschnitten als untrainierte Vergleichspersonen. Dabei ist es nicht Teil des Trainingsprogramms, die im IQ-Test verwendeten Aufgabentypen zu üben. Vielmehr wird das Arbeitsgedächtnis geschult, was offensichtlich einen positiven Effekt auch auf andere Aufgabentypen hat^{[19][20]}.

Die bei allen IQ-Tests unterstellte logische Operation zur Bestimmung von „Intelligenz“ ist die folgende: Man vermutet, dass sich eine allgemeine Fähigkeit zu intelligenten Leistungen in bestimmten messbaren Test-Leistungen äußert. Dies unterstellt eine tautologische Aufspaltung des umgangssprachlichen Begriffs Intelligenz in ein in sich selbst ruhendes Ursache-Wirkungsverhältnis: Man verdoppelt intelligente Leistungen - aus dem Bedürfnis nach Messbarkeit heraus, verallgemeinert als *abstrakte* Intelligenzleistung - in deren Äußerung (*bestimmte* messbare Test-Leistungen) und das diesen Äußerungen angeblich zugrundeliegende *Vermögen* dazu - ausgedrückt im IQ-Wert^[21]. Durch diese willkürliche Zerlegung von Intelligenz in „Fähigkeit und Äußerung derselben“ wird aus den vom Test behaupteten, sorgfältig ausgedachten *Kriterien für* das Konstrukt Intelligenz ein *Messinstrument für die* Intelligenz. Insofern trifft die berühmte Aussage, Intelligenz sei, was der IQ-Test misst, auf jede Intelligenzmessung zu und lässt somit unter der Hand überhaupt erst Intelligenz als abstrakten Gegenstand wahr werden. Bestimmte intelligente Leistungen sind inhaltlich gar nicht abstrakt vergleichbar - es sei denn, der Wille zur Auffindung abstrakter Intelligenz bringt sie genau als das, was man aus bestimmten praktischen Interessen heraus messen will, theoretisch hervor (z. B. für die Feinsteuerung von Selektion). Auf die Art und Weise des Testens kommt es im Einzelnen dann nicht an. Diese hängt weitgehend vom persönlichen Geschmack des Testers und seiner ideologischen Praxis-Orientierung („sprachfrei“, „mathematikfrei“, „milieuneutral“ usw.) ab.

Literatur [Bearbeiten]

Hans Magnus Enzensberger: *Im Irrgarten der Intelligenz. Über den getesteten Verstand und den Unverstand des Testens.* In: *Neue Zürcher Zeitung*. 11./12. November 2006, S. 28.

Stephen Jay Gould: *Der falsch vermessene Mensch*. 3. Auflage. Suhrkamp, Frankfurt am Main 1999, ISBN 3-518-28183-6.

Nicholas John MacKintosh: *IQ and Human Intelligence*. Oxford University Press, Oxford 1998, ISBN 0-19-852368-8.

Franzis Preckel, Matthias Brüll: *Intelligenztests*. Reinhardt, München 2008, ISBN 978-3-8252-3027-2.

Marc Wittmann, Andreas Eisenkolb, Christoph Perleth: *Neue Intelligenztests. Ein umfassendes Test- und Übungsprogramm*. Augustus Verlag, Augsburg 1997, ISBN 3-8043-3055-X.