

Zeitschrift: Schweizer Erziehungs-Rundschau : Organ für das öffentliche und private Bildungswesen der Schweiz = Revue suisse d'éducation : organe de l'enseignement et de l'éducation publics et privés en Suisse

Herausgeber: Verband Schweizerischer Privatschulen

Band: 35 (1962-1963)

Heft: 10

Artikel: Einige Gesichtspunkte der Anwendung und Entwicklung von Testverfahren

Autor: Tent, Lothar

DOI: <https://doi.org/10.5169/seals-851592>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 30.03.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Einige Gesichtspunkte der Anwendung und Entwicklung von Testverfahren

Dr. Lothar Tent

I. Wo können Tests eingesetzt werden?

Die Anwendung von Testverfahren in der Schulpraxis ist heutzutage nicht mehr ungewöhnlich, obwohl im deutschsprachigen Bereich viel weniger davon Gebrauch gemacht wird als im englisch-amerikanischen. In vielen Fällen, wo von Erziehern Entscheidungen verlangt werden, können Tests, richtig ausgesucht und angewendet, wertvolle Hilfsmittel sein. Sie können uns Unterlagen in die Hand geben, die es gestatten, schwierige Entscheidungen leichter zu treffen und besser zu begründen — ohne uns diese Entscheidungen selbst abzunehmen. Sie ermöglichen es uns darüber hinaus ganz allgemein, Kenntnisse, Leistungen und Verhaltensweisen von einzelnen Kindern und Jugendlichen, aber auch von ganzen Klassen oder größeren Gruppen, objektiver zu beurteilen und miteinander zu vergleichen. Diese Möglichkeiten einer objektiv-vergleichenden Leistungs-Kontrolle wird man gerade in solchen Ländern begrüßen, die über vielgestaltige Ausbildungseinrichtungen verfügen und dabei auf einheitlich straffe, zentral geregelte Lehrplanvorschriften verzichten (man vergleiche in diesem Zusammenhang die in der Schweiz üblichen pädagogischen Rekrutenprüfungen).

Ein Schwerpunkt der Zuhilfenahme von Testen in der Schule liegt auf dem Gebiet der individuellen Schullaufbahn-Beratung und -Lenkung. Das fängt mit der Frage an, ob das Kind mit Erreichen des schulpflichtigen Alters auch tatsächlich schulreif ist, oder ob man es nicht besser zunächst zurückstellt oder in anderer Weise fördert. Bei manchen Kindern ist zu entscheiden, ob sie überhaupt in der Lage sein werden, die Normalschule mit Erfolg zu durchlaufen, oder ob sie nicht in einer Sonderschule besser aufgehoben sind. Viele Kinder wollen weiterführende Schulen besuchen, und die Frage, ob und für welche sie geeignet sind, soll im voraus und möglichst zutreffend beantwortet werden. Oft ist es fraglich, ob ein Kind in die nächsthöhere Klasse versetzt werden soll oder nicht. In all diesen Fällen

wird eine Vorhersage über zukünftiges Verhalten von uns verlangt, das heißt, es muß eine Prognose gestellt werden.

Es ist klar, wie schwerwiegend solche Entscheidungen sein können. Umso weniger dürfen wir es uns erlauben, wichtige und objektive Informationsquellen, wie sie uns in modernen Testen zur Verfügung stehen, unausgeschöpft zu lassen.

Auch in den verschiedenen Berufsschulen, bei Fortbildungskursen und im Zusammenhang mit dem sogenannten zweiten Bildungsweg, wie er beispielsweise in verschiedenen Ländern der Bundesrepublik Deutschland aufgebaut wird, können Tests gute Dienste leisten. Im schulpsychologischen Dienst und in den Erziehungsberatungsstellen vieler Länder werden darüber hinaus seit langem und mit Erfolg Tests bei der Diagnosestellung an erziehungsschwierigen oder verhaltensauffälligen Kindern angewendet.

Ich habe die Bezeichnung «Test» bisher in einer allgemein-unkritischen Weise und nur im Hinblick auf die Erziehungspraxis gebraucht. Es ist bekannt, daß — abgesehen von den gemeinsprachlichen, naturwissenschaftlich-technischen und prüfstatistischen Bedeutungen des Begriffs — Testverfahren auf vielen Arbeitsgebieten der Psychologie als Mittel der praktischen Diagnostik und Eignungsvorhersage wie auch als Forschungsmittel eine große Rolle spielen. Darauf soll hier aber nicht näher eingegangen werden.

Es gibt jedenfalls inzwischen eine sehr große Anzahl sehr verschiedenartiger Testmethoden, und die internationale Literatur über Testuntersuchungen und Testmethodik ist in einer für den einzelnen Praktiker kaum noch übersehbaren Weise angewachsen. Hand in Hand mit der Verbesserung der Testverfahren geht dabei die Entwicklung der Testtheorie, der Lehre vom Aufbau und von der Analyse von Testen — eine Lehre, der in der Psychologie zweifellos eine zunehmende allgemein-methodologische Bedeutung zukommt.

II. Was besagen Testergebnisse?

Wie ist, angesichts der erwähnten Vielfalt der verschiedensten Testverfahren, der Begriff «Test» für unsere Zwecke, das heißt in der Psychologie und der empirisch-praktischen Pädagogik einheitlich zu definieren?

Ein Test ist eine standardisierte Meßmethode, die es gestattet, bestimmte Reaktionen von Individuen unter genau festgelegten, möglichst gleichen und grundsätzlich wiederholbaren Umständen zu registrieren. Insofern als die Anwendung eines Tests für jeweils ein Individuum zeitlich begrenzt ist und innerhalb dieser Begrenzung wiederum nur ein Teil aller Verhaltens- und Erlebensweisen festgehalten wird, ist der Test — seiner englischen Wortbedeutung nach — eine (Stich-)Probe, die wir aus allen vorkommenden Reaktionen des Individuums entnehmen.

Dieser Sachverhalt hat eine große Bedeutung für viele Teilprobleme der allgemeinen Testtheorie wie auch für die Folgerungen, die wir aus den Ergebnissen eines bestimmten Tests ziehen. Es wird beispielsweise von einem Test gesagt, er prüfe die Intelligenz oder messe die Begabung oder ganz bestimmte Fähigkeiten oder Eigenschaften. Dabei müssen wir uns darüber im klaren sein, daß wir damit Hilfsbegriffe verwenden, die, als Bestandteile heuristischer Modellvorstellungen, den eigentlichen Testergebnissen unterlegt werden. Die Intelligenz, die wir meinen, wenn wir sagen, A sei intelligenter als B, ist weder mit der einfachen Anzahl der gelösten Testaufgaben identisch, die wir als Ergebnis des Tests ermittelt haben, noch ist sie uns im unmittelbaren Erleben zugänglich.

Eine solche Zuordnung, nämlich die einer Fähigkeit oder Eigenschaft zu den Verhaltensreaktionen, die allein wir in einem Test tatsächlich erfassen, ist oft gar nicht so einfach und bleibt vielfach problematisch. Das mag durch ein Beispiel angedeutet werden. Ein Junge von neun Jahren habe in einem Test, der aus lauter einfachen Rechenaufgaben besteht, besser als die meisten seiner Altersgenossen abgeschnitten. Es wird gesagt, der Junge habe eine gute «Rechenfähigkeit». Ein Erwachsener habe in demselben Test genau so viele Aufgaben in der gleichen Zeit gelöst wie der Junge. Es wird gesagt, der Mann habe eine schwache oder geringe «Konzentrationsfähigkeit»; dabei wird vorausgesetzt, daß er ja «eigentlich (besser) rechnen kann».

Es wird deutlich, daß solche Aussagen zunächst nichts als Hypothesen sein können, deren Brauchbarkeit oder Plausibilität erst noch gesichert werden muß, indem wir Beziehungen zu weiteren erfassbaren Sachverhalten aufweisen. Wir werden gegebenenfalls andere oder ähnliche Tests heranziehen und andere Beobachtungen und Informationen zu verwerten versuchen.

Kein Test mißt also direkt etwas Psychisches. Die psychologischen Kategorien oder Dimensionen, in denen wir oft Testergebnisse darstellen, sind jeweils zu begründende Interpretationen, die nicht schon mit den einfachen Testergebnissen selbst gegeben sind.

Die Anwendung von gut bearbeiteten, modernen Testverfahren, die für die Schulpraxis in Frage kommen, ist jedoch weniger problematisch, als es nach den eben gemachten Ausführungen den Anschein haben könnte. Nur sollte jeder, der Testuntersuchungen durchführen will, diese grundsätzlichen Überlegungen angestellt haben und in seinen Folgerungen entsprechend zu Werke gehen. Das bewahrt uns davor, Irrtümern zu erliegen oder Behauptungen aufzustellen, deren Beweis uns schwer fallen könnte. Dabei kommt uns der Umstand zugute, daß Aussagen über Art und Wesen der untersuchten Fähigkeiten, Eigenschaften oder Persönlichkeitszüge sehr häufig gar nicht erforderlich sind und wir in der Praxis mit den numerischen Testergebnissen, so wie wir sie festgestellt haben, völlig auskommen.

So genügt es vielfach, wenn wir als Benutzer beispielsweise eines Schulreifetests wissen, daß Kinder, die in diesem Test gut abschneiden, mit großer Wahrscheinlichkeit im ersten Schuljahr erfolgreich mitarbeiten können — und umgekehrt. Wir müssen uns dabei allerdings darauf verlassen, daß diese Beziehung zwischen Testergebnis und Schulerfolg vorher vom Verfasser des Tests oder anderen Benutzern hinreichend nachgewiesen worden ist. Dabei spielt es zunächst keine Rolle, wie dieser Zusammenhang im einzelnen interpretiert werden kann (man vergleiche weiter unten den Abschnitt über die «Gültigkeit» von Testbefunden).

Damit kommen wir direkt zu Fragen der Testkonstruktion und der Testanalyse. Zuvor sei aber noch ein kurzer geschichtlicher Rückblick gestattet.

III. Aus der Geschichte der Testentwicklung

Die Geschichte des modernen Tests in Psychologie, Psychiatrie und Pädagogik ist noch keine hundert Jahre alt. Die Entwicklung von Testen oder, im Anfang, von testähnlichen Verfahren ging zu einem großen Teil von praktischen Bedürfnissen und Zielsetzungen aus. Psychologen, Mediziner, Anthropologen und Pädagogen haben zu dieser Entwicklung beigetragen. Es seien hier die Namen F. GALTON, ST. HALL, K. RIEGER, J. Mc K. CATTELL, der 1890 die Bezeichnung «mental test» einführte, H. MÜNSTERBERG, E. KRAEPELIN, H. EBBINGHAUS, TH. ZIEHEN und E. MEUMANN genannt.

Wesentliche Fortschritte wurden dann durch A. BINET und TH. SIMON (1905), durch CH. SPEARMAN und durch W. STERN gemacht, der 1912 den

Intelligenzquotienten (IQ) als einheitliches Intelligenzmaß vorschlug. Das BINETSche Staffelsystem mit Testreihen für jede Altersstufe ist in einer Reihe von Ländern mehrfach umgearbeitet und weiterentwickelt worden und wird in solchen Formen noch heute — neben neueren Verfahren — benutzt.

Im ersten Weltkrieg wurden in den Vereinigten Staaten Amerikas erstmals in großem Stil Gruppenteste angewendet (Army-Alpha- und Army-Beta-Test). Das bedeutet eine erhebliche Ökonomisierung des Verfahrens und eröffnete neue Möglichkeiten der Anwendung von Testen, so vor allem auch in Schulklassen zur Erfassung der Begabungsverteilung und zu standardisierten pädagogischen Leistungs- oder Schulfortschrittsprüfungen (educational measurement). Diese Prüfungen sind in den USA außerordentlich stark verbreitet. Dagegen hat man sich im deutschen Sprachgebiet — mit Ausnahme eines Ansatzes von O. BOBERTAG und E. HYLLA (1928) — bis zum Ende des letzten Krieges solcher Mittel kaum bedient. — Andere Techniken, wie Persönlichkeitsfragebogen (R. S. WOODWORTH 1920) und die sogenannten projektiven Verfahren (H. RORSCHACH 1921, H. A. MURRAY 1938) kamen hinzu.

Damit sind nur einige wichtige Stadien der Entwicklung angedeutet. Um den Ausbau der theoretischen Grundlagen und statistischen Techniken haben sich insbesondere CH. SPEARMAN (seit 1904), L. L. THURSTONE, J. P. GUILFORD und H. GULLIKSEN bemüht. Im englischsprachigen Bereich gibt es eine Anzahl mehr oder weniger umfassender Lehrbuchdarstellungen zur Konstruktion und Analyse von Testen. Ein vergleichbares Werk in deutscher Sprache hat bis vor kurzem gefehlt.

Diese Lücke ist erst durch das Lehrbuch des bekannten Testautors und Hamburger Professors G. A. LIENERT: «Testaufbau und Testanalyse» geschlossen worden, das in diesem Jahr im Verlag J. BELTZ in Weinheim/Bergstraße erschien (vergleiche die Literaturhinweise am Schluß dieses Aufsatzes). Nach einer Arbeit von HYLLA (1927) und neueren Teildarstellungen von H. ANGER und R. BARGMANN (1954), G. EKMAN (1955) und H. BELSER (1959) gibt LIENERT erstmals in deutscher Sprache eine zusammenfassende Darstellung über den gegenwärtigen Stand der Forschung auf diesem Gebiet. Das klar und anschaulich geschriebene Werk, auf das ich mich im folgenden weitgehend beziehe, sei dem interessierten Leser nachdrücklich empfohlen. Es ist auch bei geringen Vorkenntnissen gut zur Einarbeitung in die Methoden und Probleme der Testforschung geeignet.

IV. Welche Arten von Testverfahren gibt es?

Bevor wir auf die allgemeinen Eigenschaften eines Tests als Meßinstrument näher eingehen, wollen wir uns zunächst anhand einiger Klassifizierungsgesichtspunkte, nach denen Tests beschrieben

und geordnet werden, einen Überblick über die verschiedenen Arten von Testverfahren verschaffen.

1. Eine einfache, ohne weiteres verständliche Einteilung ist die in *Individual-* (Einzel-) und *Gruppenteste*. Gruppenteste sind in der Regel als Verfahren ökonomischer. Sie sind es auch, die in erster Linie für die Schule und für die Anwendung durch Lehrer in Frage kommen.

2. Gruppenteste sind zum weitaus größten Teil gleichzeitig sogenannte *Papier-* und *Bleistiftteste* im Unterschied zu *Materialbearbeitungs-* (oder Handlungs-)Testen, die meist als Individualtest durchgeführt werden müssen (z. B. Mosaiktest und manche Arbeitsproben).

3. Nach dem Grad der Sprachgebundenheit unterscheidet man zwischen *verbalen* und *nicht-verbalen* (z. B. auch weitgehend «kulturfrei» aufgebauten) Testen. Nicht-verbale Tests werden zur Untersuchung von Taubstummen, Sprachunkundigen und Analphabeten verwendet (der bereits genannte Army-Beta-Test war ein solcher), aber auch dann, wenn anschauliche Probleme (z. T. des technischen Verständnisses) oder abstrakte Denkaufgaben gelöst werden sollen (z. B. der Matrizen-test von L. S. PENROSE und J. C. RAVEN).

4. Ein anderer, vor allem teststatistisch wichtiger Gesichtspunkt ist die Zeit, die man für die Lösung der Testaufgaben zuläßt, und man unterscheidet zwischen *Schnelligkeitstesten* und *Niveautesten*. Bei den erstgenannten geht es darum, innerhalb einer vorgegebenen Zeit möglichst viele meist gleich leichter Aufgaben zu bearbeiten (z. B. Rechenaufgaben nach R. PAULI oder H. DÜKER und LIENERT), bei den anderen kommt es darauf an, bei großzügig bemessener oder freier Zeit und steigender Aufgabenschwierigkeit möglichst viele auch der schwierigeren Aufgaben richtig zu lösen und so das «geistige Niveau» zu erfassen (z. B. bei Wortschatztesten).

Die meisten Tests sind in dieser Hinsicht Mischformen. Auf jeden Fall ist es aber für den Testbenutzer wichtig, die in den Testanleitungen aufgeführten Zeitbegrenzungen genauestens zu beachten, da die Testergebnisse andernfalls verfälscht werden und ihre Vergleichbarkeit verlieren.

5. Nach der Art des Zustandekommens von Testergebnissen und den Möglichkeiten ihrer Deutung unterscheidet man *direkte* (Leistungsteste, Fragebogen) und *projektive* (indirekte) Testverfahren (z. B. den Formdeuteversuch nach RORSCHACH, den Thematischen Apperzeptions-Test nach MURRAY oder das Picture Frustration Study nach S. ROSENZWEIG).

Bei den projektiven Verfahren wird von der untersuchten Person erwartet, daß sie sich in das ihr

vorgelegte Material «hineinversetzt» (Projektion). Von den dabei auftretenden Reaktionen wird dann nach bestimmten Deutungsprinzipien (indirekt) auf die Merkmalsstruktur geschlossen.

Die direkten Tests sind in der Regel gleichzeitig *objektive*, während die projektiven Verfahren zu einem Teil insofern *subjektive* Methoden sind, als hier die Einordnung und Auswertung der Testreaktionen mehr oder weniger durch das subjektive Urteil des Auswerters mitbestimmt werden. Die meisten projektiven Testverfahren folgen dem Prinzip der *freien* Aufgabenbeantwortung (die untersuchte Person kann äußern, was ihr einfällt und vielfach auch soviel sie will), während die meisten objektiven Tests die *gebundene* Aufgabenbeantwortung benutzen. Hier muß die Lösung aus vorgegebenen Möglichkeiten herausgefunden werden.

6. Nach der Anzahl von Untertesten und der Anzahl der Auswertungsgesichtspunkte kann man noch grob unterscheiden zwischen *einfachen* und *komplexen* Testen. Die einfachen Tests sind in der Regel auch *eindimensional* (z. B. der Fragebogen M.M.Q. nach H. J. EYSENCK oder der Rechentest K.L.T. nach DÜKER und LIENERT oder Schulleistungsteste), die komplexen *mehrdimensional*, das heißt, es wird entweder nur ein einziges oder eine Anzahl voneinander unterscheidbarer (relativ unabhängiger) Merkmale erfaßt. Zu den zweiten gehören die projektiven Verfahren und die Testsysteme, die aus einer festen Kombination relativ selbständiger (Unter-)Tests bestehen (z. B. der Intelligenz-Struktur-Test nach R. AMTHAUER, Entwicklungsteste nach CH. BÜHLER und H. HETZER).

7. Wird ein Test, ganz gleich welcher Art, zur Eignungsvorhersage, z. B. eines Anwärter für einen bestimmten Beruf, benutzt, so spricht man von einem *Eignungstest*, wenn seine Brauchbarkeit für diesen Zweck empirisch genau überprüft ist.

8. Neben diesen mehr oder weniger formalen Unterscheidungsgesichtspunkten ist die Einteilung der Tests *nach der Art* der zu *erfassenden Merkmale* die

praktisch wichtigste. Stark vereinfachend, aber sachlich am besten gerechtfertigt, kann man zwei große, nach ihrer inhaltlichen Zielsetzung verschiedene Gruppen unterscheiden: die *Leistungsteste* und die *Persönlichkeitstests* (oder charakterologischen) *Tests*:

Leistungsteste:

- a. *sensorische*, die Sinnestüchtigkeit untersuchende Tests (z. B. Farbtafeln nach J. STILLING)
- b. *motorische* Leistungsteste (z. B. Punktieren, Perlenaufzählen)
- c. *Konzentrations-Leistungsteste* (z. B. Rechenteste, Sortierungsaufgaben)
- d. *Wissens- oder Kenntnisteste* (z. B. Schulfortschritts- oder Schulleistungsteste, etwa: «Rechtschreibtest für das achte Schuljahr», Zahlenrechnen für das vierte Schuljahr», Herausgeber: Hochschule für Internationale Pädagogische Forschung, Frankfurt)
- e. *spezielle* Intelligenz-, Entwicklungs- und Begabungsteste (z. B. Gedächtnisteste, Schulreifeteste, standardisierte Arbeitsproben)
- f. *allgemeine* Intelligenz-, Entwicklungs- und Begabungsteste Testsysteme wie z. B. die von BINET, H. BIÄSCH, R. MEILI, D. WECHSLER, HETZER, AMTHAUER, W. HORN)

Persönlichkeitsteste:

- a. Tests zur Erfassung von *Interessen, Einstellungen* und Werthaltungen (z. B. der Persönlichkeits- und Interessentest P.I.T. nach E. MITTENECKER und W. TOMAN, die Einstellungsskalen von THURSTONE und L. W. FERGUSON)
- b. Tests zur Erfassung von *Persönlichkeitszügen* und der Typenzugehörigkeit z. B. Persönlichkeitsfragebogen nach R. G. BERNREUTER, R. B. CATTELL, EYSENCK)
- c. Tests zur Erfassung *klinischer* Persönlichkeitsveränderungen (z. B. Beobachtungsbogen nach J. R. WITTENBORN, Gestalterfassungstest nach L. BENDER)
- d. Tests, mit denen man die gesamte *Affekt- und Persönlichkeitsstruktur* zu erfassen beabsichtigt (z. B. Formdeutversuche, thematische Apperzeptionsteste, Farbteste). Diese Tests stehen erst am Anfang ihrer Entwicklungsmöglichkeiten.

Dieser Überblick läßt einige der Aufbauprinzipien und die vielseitigen Anwendungsmöglichkeiten der modernen Testverfahren erkennen. Im nächsten Abschnitt sind nun die wichtigsten Eigenschaften darzustellen, die alle Tests — ganz gleich welcher Art — besitzen sollten.

(Schluß folgt)

Education stimulante, facteur de santé mentale

par F. Hotyat

Enseignants, nous nous proclamons volontiers être en même temps des éducateurs. Qu'est-ce à dire? Que, vivant parmi les enfants et les adolescents, auréolés du prestige de nos fonctions, nous induisons en eux des comportements, des réactions affectives, des modes de pensée qui, répétés au fil des années,

marquent véritablement les traits de leur personnalité.

Ainsi, nous estimons qu'il est en notre pouvoir — certes dans une mesure relative — d'aider nos élèves, à devenir des adultes confiants dans la vie, s'acceptant eux-mêmes et acceptant autrui, doués d'assez