

Leistungsgrenzen des Tests

Von Wilhelm Salber, Bonn

Das Wort „*Testen*“ beginnt langsam in der deutschen Sprache Lehnwort zu werden. Schon ist die Verwendung nicht mehr auf die Fachterminologie beschränkt; illustrierte Zeitschriften und Zeitungen haben für weitreichende Verbreitung gesorgt.

Dabei herrschen über die Leistungsmöglichkeiten die widersprechendsten Meinungen. Der vorliegende Beitrag soll ein Versuch sein, sich gerade mit dieser Frage auseinanderzusetzen.

Um die Leistung und die Leistungsgrenzen des Tests darstellen zu können, ist es zweckmäßig, von den wesentlichen Merkmalen und von der besonderen Eigenart auszugehen. An Hand der verschiedenen aufzuweisenden Momente läßt sich über das Wesen des Tests und seine Leistung am einfachsten eine Aussage machen.

1. Der psychologische Test ist ein Kind des psychologischen Experiments. Mit ihm teilt jeder Test das Merkmal der *Künstlichkeit*. In einer provozierten Situation wird der Prüfling vor bestimmte Aufgaben gestellt, auf die er reagieren soll. Die Künstlichkeit hat den Vorteil, daß der Untersuchende geeignete einfache und durchschaubare Handlungsabläufe auswählen kann, daß er die Bedingungen bzw. die Besonderheiten besser übersieht, daß er normierte Vergleichsmöglichkeiten hat und nicht dauernd auf einen zufällig günstigen Moment zur Erfassung der ihn interessierenden Sachverhalte warten muß.

Nachteile der Künstlichkeit liegen in der notwendigen Isolation dessen, was sonst in einem lebendigen Zusammenhang ist, in der Ausschaltung der Spontaneität; zudem erschwert das Odium der Prüfungssituation oft die Untersuchungen.

2. Ein zweites wesentliches Moment des Tests zeigt sich in *Sachcharakter* und *Sachlichkeitscharakter* von Test und Testergebnis. Tests sind Sachen einmal in der mehr

oberflächlichen Bedeutung, daß sie angefaßt und transportiert werden können; sie können von verschiedenen Menschen benutzt werden usw. Darüber hinaus bedeutet Sache hier „Bestehendes“ und „Gegebenes“, in bestimmter Weise geformtes *Ding*. Ding, das heißt in einem allgemeinsten Sinne „Nicht-Menschlich“ und bedeutet, daß in der Sache „Test“ — beim Test und beim Testergebnis — die Subjektivität des Untersuchenden ausgeschaltet ist. (Das gilt freilich nicht vom Vorgang des Testens, allerdings ist damit noch kein psychischer Befund vorhanden. Es muß aber berücksichtigt werden, welche Qualitäten den Gegebenheiten der Sache selbst zukommen.)

Weiterhin bestimmt sich der Sachcharakter des Tests als Zeugcharakter im Sinne *Heideggers*. Der Test ist ein Erzeugnis zu einem Zweck. Das wesentliche Merkmal, das *Heidegger* dem Zeug zuschreibt, die Verlässlichkeit und als deren Folge die Dienlichkeit, soll auch einen besonderen Vorzug der Sache Test ausmachen.

Daß das „Nicht-Menschliche“ der Sache allerdings bisweilen sogar zum Unmenschlichen werden kann, zeigt sich bei Reaktionen einzelner Vpn., die gegenüber den für sie unangenehmen Tests mit Schock „versagen“.

Darüber hinaus birgt der Sachcharakter von Test und Testergebnis noch eine weitere Einschränkung: Der Test vermag überhaupt nie von sich aus zu erfassen, was nun das Individuelle des Lebendigen als Individuelles bestimmt. Eine Sache ist und bleibt unlebendig: niemals kann von einer Testsache aus das — wie *Wellek* es nennt — Mysterium der Person entschleiert werden. Ihr fehlt Plastizität und Gestaltungskraft, ihr fehlt Einfühlungsfähigkeit und die Möglichkeit, Wesenhaftes von sich aus herauszuspüren. Wenn die Sache falsch angesetzt wird (s. u. Medialität), kann es leicht zu „Unsachlichkeiten“ kommen. Das rein quantitativ-sachliche Abmessen des Testergebnisses hört bei den charakterologischen Momenten auf. Genaue Zahlen im Test haben ihre Grenze in der qualitativen Bestimmtheit des Seelischen.

Das Moment des Sachcharakters birgt in sich für den Test auch einen Wertanspruch: Sachlichkeit in der Bedeutung von Objektivität besagt für einen psychologischen Test, daß er in das System erprobter Erfassungsweisen des Seelischen aufgenommen ist. Jede Äußerung eines Menschen auf irgendeine Sache, sei es in einer künstlichen, sei es in einer natürlichen Situation, zeigt Psychisches. Insofern ist eben alles Test. Aber um den Anspruch „psychologisch verwertbarer Test“ zu erfüllen, muß noch etwas mehr vorhanden sein. Der Test muß geeicht, vergleichbar und systematisiert, seine Ergebnisse durch wissenschaftliche Interpretation für stichhaltig befunden sein.

Ein Zug, der dem Test anhaftet, weil er eine Sache ist, ist das Fertigsein. Wenn der Test feststeht, die Ergebnisse systematisiert sind, dann ist er auch in den Zustand des „Fertigen“ übergegangen. Das bedeutet Sicherheit und Garantie für Zuverlässigkeit, aber es birgt auch die Gefahr des Bekanntwerdens und des Überholtwerdens von neuen Erkenntnissen. Wie *Stern* gezeigt hat, werden von der Theorie her die Testergebnisse neu interpretiert, eventuell ganze Testbatterien über den Haufen geworfen. (Beispiele für das Bekanntwerden und entsprechende Auswirkungen liegen im *Rorschach* vor allem bei den Farb- und Sexualdeutungen; auch Veröffentlichungen des Schneeballbildes von *Binet* in der *Zeitung* führten zu verfälschten Resultaten bei der Testuntersuchung).

Besonders gefährlich kann das werden, wenn das Fertigsein der Sache Test auch mit einer „fertigen“ Terminologie der Auswertung verknüpft ist. Dann kommt es

u. U. dazu, daß in die eigenen psychologischen Anschauungen des Auswertenden eine Terminologie einbricht, die mit dem gleichen Wort etwas anderes meint. So z. B., wenn kritiklos die Affektivität im Sinne *Rorschachs-Bleulers* in einer Terminologie gebraucht wird, die der von *Lersch* nahesteht.

Von der Art der „Sachen“ her hat man unterschieden: verbale und nicht verbale Tests, — in einer anderen Hinsicht: Gestaltungs- und Projektionstests sowie Leistungstests.

Aber mit diesen Unterscheidungen wird auf den zentralen Anspruch des Tests nicht eingegangen: auf den Anspruch, daß Psychisches an diesen Sachen sichtbar wird. Zum Sachcharakter gehört nämlich als wesentliche Bestimmung — das ist der Witz des Testens —, daß sich an den Sachen Veränderungen zeigen. Die Veränderungen, die an den Tests erfolgen, müssen durchaus nicht immer durch manuelle Handlungen hervorgerufen werden. Veränderung bedeutet vielmehr nur: die Sache hat nicht mehr die „Unberührtheit“, die gegenüber der Vp. bestand, ehe die provozierte Testsituation eintrat. Das ist der Fall, wenn die Vp. einen undurchstrichenen Ausstreichtest durch Ausstreichen verändert, wenn ein Zeichen des *Wartegg* zu einer Zeichnung wird, wenn ein Tintenklecksbild als Nikolaus gedeutet wird. Zwischen der Instruktion über die Sache und dem Endergebnis hat ein Prozeß stattgefunden, der die Veränderung bewirkt.

Für die Umgrenzung dessen, was ein Test ist, ist die Bestimmung sehr wichtig, daß das Endergebnis des Prozesses ebenfalls einen sachlichen Niederschlag findet; das psychische Werden, der Prozeß, wird nur gefaßt, wenn ein „sachliches“ Ergebnis darauf hinweist. Hierdurch unterscheidet sich der Test vor allem von der Verhaltensbeobachtung. Meist ist der Sachcharakter des Testergebnisses in eine Formel zu bringen. Der Anspruch des Ergebnisses auf Sachcharakter ist die Grundlage jeder Blinddiagnose.

Die Sache Test merkt gleichsam Psychisches an, sie weist Veränderungen auf wie ein Plastilinklumpen, der den Fingereindruck registriert. Und da die Veränderung an einer Sache fixierbar ist, läßt sich von hier aus auch der Exaktheitsanspruch verstehen.

Die feststellbaren Veränderungen sind die Testergebnisse und auf sie kommt es eigentlich an. Die Frage nach der Leistung des Tests wäre ganz wörtlich genommen damit zu beantworten: er zeitigt die Testergebnisse. (Von diesen aus müßte, wenn man genau sein wollte, die Frage nach den Leistungsgrenzen gestellt werden.) Daher ist eine Unterteilung der Tests am besten von den Hinweisen der Testergebnisse auf psychologische Einteilungsprinzipien vorzunehmen wie sie *Thomae* in Fähigkeits-, strukturellen und thematischen Gruppierungen gegeben hat; denn damit ist eine tiefgreifende Unterscheidung von den Ansprüchen des Tests her aufgewiesen.

3. Die Bedeutung der Veränderung von der Testsache zum Testergebnis weist auf ein weiteres Merkmal des Tests hin, das als *Medialität* des Tests bezeichnet werden kann. Wenn man vom Verlauf einer Testuntersuchung ausgeht, gelingt es am ehesten, sich diesen Zug klarzumachen: Versuchsperson — Test — Testergebnis (Formel) — Versuchsleiter — Interpretation. (Beispiel: Karl X. — Rorschach Tafel I — Fledermaus (GF TV) — Interpretation nur aus Gesamt der Lösungen möglich).

Der Mediumcharakter des Tests stellt einmal heraus, daß der Vl. die Reaktionen des zu Untersuchenden in einem Medium, einer dazwischen geschalteten Sache, auf-

fängt. Durch den Begriff der „Medialität“ wird die Spiegelung der Handlungen in der Sache als einem *Mittel* besonders anvisiert. Der Test läßt zugleich aber eine Umformung der Aussagen auf eine Interpretationsmöglichkeit hin vollziehen. Er ist gleichsam ein *Transformator*, der statt der möglichen Fülle von nicht sofort rubrizierbaren Äußerungen der Vp. ganz bestimmte „faßbare“ Äußerungsformen ermöglicht.

Zwischen den zu untersuchenden Menschen und dem von ihm festzustellenden psychischen Befund sind so eine Reihe von Zwischengliedern eingeschaltet. Die „Sache“ Test steht einmal zwischen Vp. und Vl.: d. h. für den Untersuchten, daß der Test ihm so angepaßt sein muß, daß er ihm in der Untersuchung überhaupt als ein Medium dienen kann. Denn das sachliche Medium, das objektiv registrieren soll, transformiert verfälschend, wenn ein Test dem einzelnen Menschen nicht angepaßt ist, sei es in altersmäßiger, soziologischer oder nationaler Beziehung. Hier ist die Grenze jeder vermittelnden Sache deutlich. Außerdem können auch Ermüdung oder z. B. mangelndes Interesse Momente sein, die das Medium nicht eine rechte Wirkung haben lassen. Wenn diese Dinge nicht berücksichtigt werden, kann sich als Resultat vielleicht ergeben: Vp. = 0.

Der andere Pol des Mediums „Test“ ist der Vl. Er hat die Aufgabe, die sachlichen Ergebnisse umzusetzen in das Bild einer psychischen Struktur. Dazu ist für ihn die Kenntnis psychischer Gesetze, Strukturprinzipien usw. notwendig. — Die Bedeutung der Haltung des Vl. darf nicht unterschätzt werden.

Die Testergebnisse stehen zwischen Test und Interpretation. Diese „Medialität“ sucht eine Hilfe zu geben für ein Hauptproblem der psychologischen Diagnostik: das Problem der psychologischen Kategorien. Denn bestimmten Testergebnissen sind oft bestimmte psychologische Kategorien zugeordnet. Von der Handlung der Vp. auf eine Sache hin fließt gleichsam eine erformte Verbindung zu einer psychologischen Erkenntnis. Hier können Fehlerquellen liegen, wenn nicht genügend geklärt ist, ob die Grundannahmen der Testautoren richtig sind, welche Terminologie zugrunde liegt und was der Test prüft.

Für den ganzen Bereich der Medialität gilt ein Gesetz, das auf keinen Fall vernachlässigt werden darf: das Gesetz der sachlichen Identifikation. Die Übereinstimmung des in der Wirklichkeit gegebenen Sachverhalts mit dem Gegenstande, der untersucht werden soll und der zur Darstellung (auch zu logisch-begrifflicher Darstellung) gelangt, muß überall gewahrt sein (*Ach*).

4. Die Medialität und der Sachcharakter hängen noch mit einem andern Moment des Tests zusammen: dem *Rastercharakter*. Der Begriff des Rasters sucht auf die innere Struktur der Medialität einzugehen, er ist die „mittelnde Sache“ in sich selbst betrachtet.

Der Test stellt ein Gitter dar, — einen Raster — der einiges auffängt, anderes nicht erreicht oder von ihm nur gestreift wird. Ist der Raster zu differenziert, nähert er sich der Situation des Lebens und gibt die ganze Vielfalt wieder. Das ist jedoch nicht die Intention des Untersuchers. Wie aus dem ungeheuren Bereich der Erscheinungen Gesetze gefunden werden sollen, so soll auch im Test-Raster nur Wesentliches aufgefangen werden. Der Test versucht Klarheit und Sonderung zu erreichen.

Der Test soll also auffangen, aber ein einzelner Test kann nicht alles erfassen, obschon es im Menschen da ist. Aus dem Begriff des Rastercharakters folgt, daß, um

bestimmte Reaktionen zu erfassen, andere psychische Momente vernachlässigt werden müssen.

Je prägnanter ein Testergebnis fassen soll, je spezieller ein Testergebnis für das stehen soll, was es zu fassen sucht, desto eher besteht die Möglichkeit des genauen Ergebnisses in einem Punkt, während andere Gegebenheiten nicht erfaßt werden. Ein Ergograph-Ergebnis vermag nichts über die geistige Produktivität auszusagen, das Intelligenzalter nichts über die Thematik. Dem entspricht ein zweites Moment: je engumrissener die Frage des Tests ist, um so geringer ist die Gewähr, daß diese Frage gerade differenzierte Ergebnisse bei allen Individuen zeitigt. Es besteht das Verhältnis der umgekehrten Proportionalität von Prägnanz und Erfassungsbereich.

Damit hängt die größere oder geringere Komplexität der Testergebnisse zusammen. Je weiter der Bereich ist, innerhalb dessen Antworten der Vp. ausgelöst werden können, um so eher gibt der Test ein „Zeichen“, aber um so unprägnanter und weniger speziell gefiltert ist es auch. Denn viele und vieles können darunter gefaßt werden. Dennoch kann auch im Falle einer großen Komplexität eine *sichere* Aussage möglich sein; nur wird dabei das Individuelle nicht genügend erfaßt. (B—antwort im *Rorschach*, „Gestaltung“ im *Ruppschen* Wabentest.) Die Sicherheit leidet aber sofort dann, wenn aus Komplexem Spezielles erdeutet werden soll.

Außer vom Raster des Tests kann man auch von einem Raster der Ergebnisse sprechen. Dieser Raster der Ergebnisse ist nicht in seinen Lücken durch immer neue Rasterpunkte ausfüllbar. Hier ist eine Grenze gegeben. Die im Raster aufgefangenen Momente dürfen nämlich nicht als Elemente angesehen werden, die nur zusammengesetzt zu werden brauchen. Eine Aneinanderreihung von Testergebnissen ist unmöglich. Denn in allem äußert sich die gesamte Persönlichkeit. Alles, was wir im Test finden, sind nur unselbständige Momente der lebendigen, sinnvoll strukturierten Einheit (*Stern*). — Der Mensch ist kein Mosaik, jede Einzelheit erhält ihre Bedeutung erst vom Ganzen: nicht Elemente sind gegeben, sondern nur Anklänge an Umfassendes. Testuntersuchungen legen Schnitte durch die Gesamtpersönlichkeit. (*Pintschovius, Stern*.) Die Verbindung zum Ganzen kann jedoch nicht mehr vom Test geleistet werden. Sie ist Sache der Interpretation.

5. Die besondere Struktur des Tests (Raster), die das Wesentliche für die jeweilige Fragestellung herauszuheben sucht, muß auch deswegen besonders beachtet werden, weil ein weiterer Wesenszug des Tests in seiner Statik liegt. Das *Statische* des Tests verdeutlicht zugleich einen besonderen Anspruch dieser diagnostischen Methode, den Anspruch nämlich, in einer konzentrierten, genau durchdachten Weise gleichsam in einem Augenblick und an einem „Punkt“ etwas zu erfassen, was sonst u. U. nur in einer genauen Beobachtung über ein Leben erfaßt werden kann: aus der „mikroskopischen“ Untersuchung (*Roemer, Thomae*) soll ein Schluß auf die Gesamtdynamik möglich werden.

Weiterhin: neben der Statik der „Sache“ Test ist ja besonders das statische Endergebnis wichtig, das einen „kompakten“ Hinweis auf überdauernde Strukturen geben will. Statt einer dauernden Verhaltensbeobachtung über längere Zeit wird versucht, mit dem „konzentrierten“ und „konzentrierenden“ Mittel des Tests ein wesentliches Ergebnis zu erzielen. Indem aber ein Moment = Augenblick ausreichen soll, viele Momente und das diese Momente Begründende zu erfassen, spielt die Augenblicksbedingtheit natürlich eine große Rolle. Tests, die eine hohe Variabilität

aufweisen, sind nur brauchbar (z. B. der *Pfister-Test*), wenn durch mehrere Proben die Bewegungsgrenze zu fixieren versucht wird.

Schließlich ist auch das Ziel relativ „statisch“. Im Test soll ein Zustand erfaßt werden (*Meili*), u. U. ein Gefüge von Dispositionen, ein Charakter in seiner Gegenwärtigkeit. Wenn die Frage gestellt wird: wie ist die geformte Psyche in der Gegenwart, wäre es falsch, sie nur von der Vergangenheit, z. B. von einem Trauma her zu sehen. Gegenüber dem *einen* Punkt, der in einem Kindheitserlebnis gefaßt wird, soll in einem psychologischen Befund der *Charakter* in seinem Aufbau, seiner Struktur, seiner Energieverteilung, seiner besonderen Ausprägtheit erfaßt werden. Es soll z. B. nicht allein der Grund des So-Seins und auch nicht allein das Ziel des Menschen aufgewiesen werden; sondern das Eingebautsein von Gründen und Zielen in eine Persönlichkeit. Unter dem Aspekt des Wandels der Persönlichkeit betrachtet, wird damit noch eine weitere Begrenzung des Tests sichtbar: eine einmal durchgeführte Testuntersuchung kann für einen Menschen nicht einen für sein ganzes Leben repräsentativen Befund liefern.

6. Ein letztes Moment des Tests, das die übrigen Charakterisierungen ergänzt und übergreift, ist der *Zeichenscharakter* der Testergebnisse. Hier liegt das psychologische Schwergewicht der Testmethode; ohne Würdigung dieses wesentlichen Tatbestandes ist der Hinweis auf die übrigen Momente eine unvollständige Erkenntnis. — Tests zeitigen Zeichen: *aliquid stat pro aliquo*.

Auch hieraus ergeben sich wieder einige Einsichten in Leistungen und Leistungsgrenzen des Tests.

Testergebnisse müssen zugeordnet werden. *Stern* weist darauf hin, daß das Testergebnis unmittelbar nur etwas über die Oberflächenschicht aussagt. Da psychische Eigenschaften nicht direkt gesehen werden können, muß *gedeutet* werden. Die geringsten Deutungsarbeiten erfordern die oberflächenhaften Leistungen. Hier erfolgt eine einfache feststellende Bewertung mit Plus oder Minus und entsprechenden Zwischenwerten oder — wie z. B. bei *Binet* — eine psychologische Deutung auf Grund prozentualer Berechnungen. Je näher die Deutung der reinen Leistungssituation kommt, je eher ist die Gewähr gegeben, daß keine Täuschung vorliegt, je weniger aber ist eine Erdeutung hintergründiger Wesenseigenschaften möglich. Mit steigender Unspezifität und steigender Entfernung von der Leistungstatsache wird die Deutung schwieriger — und psychologisch bedeutsamer (ohne notwendigerweise unsicherer zu werden).

Testergebnisse sind vieldeutig. Eine Zuordnung je eines einzelnen Testergebnisses zu je einer einzelnen Fähigkeit, Disposition, Eigenschaft ist nicht möglich. Hinter ähnlichem Verhalten kann auch im Test verschiedenes als Begründung liegen. Gerade hierauf hat *Wellek* bei einer Besprechung der Signierungsschwierigkeiten im *Rorschach* erst vor kurzem eindringlich hingewiesen.

Aus der Komplexität der Testergebnisse und der Konzentriertheit des Tests ergibt sich zusammen mit der Notwendigkeit zu deuten: die Notwendigkeit der Interpretation. Das Zusammenbringen der Einzelbefunde in ein Gesamtbild setzt Wissen um das voraus, was Struktur, Ganzheit, Persönlichkeit ist, wie verschiedene psychische Momente aufeinander wirken usw. Die Neuauflage des Lehrbuchs der Diagnostik von *Meili* verzichtet bewußt auf den „schwierigsten Teil der Diagnostik“, nämlich „die Verarbeitung der Untersuchungsbefunde über eine Person zu einem

einheitlichen Persönlichkeitsgutachten“, denn die Ausdeutung der Testresultate sei so komplex, daß sie Vertrautheit mit Tests, allgemeine psychologische Ausbildung und viel praktische Erfahrung schon voraussetze.

Ob nun hier der erforderliche Vollzug als „Brückenschlag des Verstandes“ (*Pintschovius*) oder als Intuition (*Wellek*) bestimmt wird, immer erweist sich etwas für die Verwertung des Tests notwendig, das Analogien zur künstlerischen Gestaltung aufweist. Die letzte Sicherheit liegt in der darauf folgenden inneren Evidenz (*Pintschovius*).

Zusammenfassung: Die Tests sind sachliche Zwischenglieder der psychologischen Untersuchung, sie stellen ausgewählte Situationen dar und ermöglichen — trotz der Kürze der Untersuchung — verlässliche Hinweise für psychische Befunde.

Aber die Testuntersuchungen können nur Anzeichen liefern, die Material zur Verfügung stellen. Jedes der aufgewiesenen Momente (Künstlichkeit, Sachcharakter, Medialität, Rastercharakter, Moment des Statischen und Zeichencharakter) weist auf die Grenzen des Tests hin und selbst für die schönste Testbatterie gilt: „Das Feinste fällt durchs Sieb.“ (*W. Busch*.)

Literatur

- Meili, Richard:* „Lehrbuch der Psychologischen Diagnostik.“ Huber, Bern 1951.
Pintschovius, Karl: „Die Psychologische Diagnose“. Lehmann / München-Berlin 1940.
Stern, William: „Persönlichkeitsforschung und Testmethode“. Jahrbuch d. Charakt. VI. Jg. 1929. Pan-Verlag, Berlin.
Thomae, Hans: „Experimentelle psychologische Diagnostik.“ Fortschritte d. Neurologie. Psychiatrie. 19. Jg. 1951.
Wellek, Albert: „Das Experiment in der Psychologie.“ Studium Generale 1952.