

lich des Entwurfs der Items und von Vorversuchen an Vpn. mit dem Ziel, die Test-*Vorform*, die *Präliminar-Form*, zu entwickeln. In der 2. Phase erfolgt die  $\hat{I}$  Itemselektion der Test-Vorform, die Festlegung der Test-Endform oder *Final-Form*, durch Auswahl geeigneter Items auf Grund von aus einer Stichprobe gewonnenen Kennwerten und die Überprüfung der Gütekriterien ( $\hat{I}$  Reliabilität,  $\hat{I}$  Validität). In der 3. Phase wird die Endform an einer neuen Stichprobe derselben Population überprüft, in der 4. Phase schließlich geeicht, d. h., an größeren repräsentativen Stichproben derselben oder einer in bezug auf Alter, Bildungsgrad oder Herkunftsorte ähnlichen Population werden Normen gewonnen. In allen Phasen der T. sind dabei folgende Aspekte zu beachten:

- 1) die zu diagnostizierende Person-Eigenschaft (Dimension, Disposition, Indikand) entweder habituellen Art wie Fähigkeit, Kenntnisstand, Temperament, Interessen, Haltungen oder aktueller Art wie Stimmung, Grad der Ermüdung oder Grad eines Rausches;
  - 2) die damit verbundene Validierung des Verfahrens, d. h. Beantwortung der Frage, wie weit es die zu diagnostizierende Eigenschaft tatsächlich mißt;
  - 3) Auswahl der Stichprobe(n) und damit der Population<sup>^</sup>, auf die der Test anzuwenden bzw. in denen die Itemanalyse durchzuführen ist;
  - 4) das Medium, z. B. schriftliche Vorlagen, Bilder, Spielmaterial, Film oder Tonband;
  - 5) die Einwirkung durch den Untersucher, z. B. durch Instruktionen, durch Aufgabenstellungen, Fragen, Erläuterungen, oder auch Gestaltung der Situation;
  - 6) Situationsbedingunge $\hat{u}$  i. w. S., z. B. Einzel- oder Gruppensituation, Gestaltung der Umgebung, z. B. ihre Störfreiheit, Tageszeiten, Niveau- oder Geschwindigkeitstest;
  - 7) relevantes Verhalten bzw. Reaktion des Untersuchten, d. h. diejenigen Aspekte des Verhaltens, aus denen Aussagen über die zu diagnostizierende Eigenschaft abzuleiten sind;
  - 8) Skalierung der relevanten Verhaltensweisen zwecks Protokollierung, um Daten zu erhalten; bei Aufgaben z. B. muß feststehen, was als falsch/richtig oder falsch/teilweise richtig/richtig gilt, bei Fragen z. B. die Antworten ja/nein oder ja/unentschieden/nein;
  - 9) Zusammenfassung der Ergebnisse der Items zu Gesamt-Werten in Form von Rohwerten oder Rohpunktwerten, z. B. durch Angabe der Anzahl der richtigen Lösungen oder von Bejahungen;
  - 10) Formulierung des Ergebnisses, zumeist durch eine Normen-Skala, z. B. durch Prozentrang-, Standardwert- oder Äquivalent-Normen, oder auch Erfüllung eines bestimmten Minimal-Kriteriums, jeweils in bezug auf eine oder mehrere Eigenschafts-Dimensionen.
- Diese 10 Aspekte sind bereits bei der Planung zu

berücksichtigen; in den weiteren Phasen erfolgen dann weitere Festlegungen bzw. Präzisierungen. Um genügend zuverlässige Kennwerte für die Verfahrensoptimierung zu erhalten, ist es notwendig, die empirische Überprüfung der Vorform und der Endform an mindestens dreistelligen Zahlen von Vpn. vorzunehmen; die Eichung erfolgt i. allg. an mehreren Tausend Vpn., die bestimmte Populationen repräsentieren.

In jeder Phase werden Effektivitätskontrollen vorgenommen. In den Vorversuchen wird i. allg. die Praktikabilität des Verfahrens überprüft, und es werden Festlegungen in bezug auf die Aspekte 1) bis 8) getroffen. Mit der Vorform ist die Methode weitgehend festgelegt; spätere Veränderungen betreffen fast nur Vereinfachungen, das Fortlassen und eventuell Umstellungen von Items. Bei der  $\hat{I}$  Itemanalyse wird insbesondere anhand von Schwierigkeitsindizes bei Tests bzw. Popularitätsindizes bei Fragebogen und von Diskriminationsindizes, d. h. Trennschärfe- bzw. Validitätsindizes, oder aber bei heterogenen Itemreihen mittels Faktorenanalyse eine Itemselektion vorgenommen, deren Resultat die Endform ist. Hier und in den weiteren Phasen dienen die statistischen Gütekriterien zur Effektivitätskontrolle; es lassen sich vorläufige oder endgültige Normen angeben.

Da Normen in Abhängigkeit von gesellschaftlichen Veränderungen, z. B. vom betreffenden Bildungssystem, und auf Grund neuer theoretischer Erkenntnisse veralten können, wird die Überarbeitung von Tests und Fragebogen, die sog. Test-Revision, notwendig; bei der Anwendung auf neue Populationen, insbesondere bei Übertragungen von verbalen Methoden in andere Sprachen, wird jeweils eine Adaptierung erforderlich. Sowohl Revisionen als auch Adaptationen betreffen i. allg. alle Phasen der T.

Da standardisierte Verfahren mittels geeigneter Lesegeräte maschinell oder elektronisch auswertbar sind und z. T. nicht die Anwesenheit eines Untersuchers erfordern, können in der *Computerdiagnostik* weitgehend die Mittel der EDV zur Auswertung und auch Untersuchung eingesetzt werden.

**Testhypothese:** Erwartungsbildung über die Konsequenz eines Entscheidungsschritts im Begriffsbildungsprozeß, die durch den Vergleich mit dem tatsächlichen Prüfungsangang der entsprechenden Objektordnung zum Informationsgewinn führt. Ein spezieller Fall, nämlich die Begriffshypothese, liegt dann vor, wenn die mutmaßliche Begriffsstruktur den Inhalt der T. bestimmt.

**Testinvarianz** \ logistisches Testmodell.

**Testkontinuum**  $\hat{I}$  Testtheorie, psychologische.

**Testmodell, logistisches:** spezieller Ansatz der psychologischen  $\hat{I}$  Testtheorie, der das Antwortverhalten von Meßobjekten in dichotomen Meßvariablen, die nur zwei alternative Meßergebnisse 0 und 1 zulassen, analysiert und beschreibt. Gleich-