

tensanalyse i. allg. keine direkten Meßverfahren existieren, sondern über indirekte Beobachtungsdaten ein Rückschluß auf den Ursachenkomplex des jeweiligen Verhaltensphänomens durchgeführt wird. Damit ist impliziert, daß eine sinnvolle T. voraussetzt, daß die Beobachtungsdaten Erscheinungen sind, die auf wesentlichen Zusammenhängen des Phänomenbereiches basieren. Sie gilt es zu erfassen.

Da die Beobachtungsdaten i. allg. zufallsabhängig sind, folgt, daß der Schluß von den zufälligen Beobachtungen auf notwendige Gesetzmäßigkeiten eine Modellmethodik erfordert, die auf statistischen Grundlagen aufgebaut ist. Die Gesetzmäßigkeiten gestatten die Erfassung qualitativer bzw. quantitativer Meßdaten X_{jj} als Beobachtungsergebnisse, die von einem Meßobjekt oj , z. B. einer Vp. oder deren Eigenschaften, durch eine Meßvariable x_{-} , z. B. eine Testaufgabe, erzeugt werden und damit von Oj und x_{-} abhängen, d. h., $x_{-j} = f(oj, x_{-})$. In diesem Zusammenhang wird auf der modelltheoretischen Ebene jede Meßvariable als eine statistische Zufallsgröße angesehen und das einzelne Beobachtungsergebnis als eine Realisierung dieser Zufallsgröße. Ein modelltheoretischer Ansatz der T. gibt dann Auskunft über die Ursachenkomplexe und Wirkfaktoren dieser Zufallsgröße und nutzt dazu im wesentlichen ihre statistischen Kenngrößen und Verteilungsparameter aus.

Die Spezifikation dieses statistischen Zusammenhanges ergibt die unterschiedlichen Modellansätze der p. T., von denen gefordert wird, daß aus ihnen nach Abschätzung der Modellparameter durch empirische Daten Aussagen über die Meßobjekte, über die Meßvariablen oder über beides abgeleitet werden können, die diese für den Phänomenbereich kennzeichnen (f Itemparameter). Dabei müssen sowohl die Meßvariablen durch Kennwerte, z. B. durch Itemschwierigkeit, Trennschärfe, Reliabilität, Validität, Homogenität, als auch die Meßobjekte durch Kennwerte, z. B. Fähigkeitswerte, so beschrieben werden, daß aus der Wirkung dieser Parameterwerte die beobachteten Testpunktwerte erklärbar und ableitbar werden. Gleichzeitig soll es möglich sein, aus diesen Überlegungen Aussagen über die Adäquatheit des Modells für den untersuchten Gegenstandsbereich abzuleiten.

Einen der ältesten Modellansätze der p. T., der auch heute noch weit verbreitet ist, stellt die sog. klassische Testtheorie dar (GULLIKSEN, GUILFORD, SPEARMAN, LIENERT). Sie spezifiziert den obigen Zusammenhang zwischen Meßobjekt oj und Meßvariablen x_{-} durch die beiden folgenden Postulate:

1) Für jedes Meßobjekt o_{-} existiert bei der Messung mit der Meßvariablen x_{-} ein wahrer Meßwert t_{jj} , der als Erwartungswert $t_{jj} = E(x_{-j})$ über unabhängigen Meßwiederholungen definiert wird. Das bedeutet, daß in jedem Beobachtungsergebnis X_{jj} ein wesent-

licher, systematischer Anteil enthalten ist, der als wahrer Meßwert bezeichnet wird.

2) Jede Messung enthält einen Meßfehler e_{jj} , der bei der Messung von Oj durch x_{-} zufällig entsteht. Dann setzt sich das beobachtete Meßergebnis $X_{jj} = t_{jj} + e_{jj}$ aus dem wahren Meßwert t_{jj} und dem zufälligen Meßfehler e_{jj} zusammen.

Aus den Beobachtungsdaten X_{jj} über einer Stichprobe von Meßobjekten lassen sich die wahren Testpunktwerte t_{jj} z. B. durch eine \hat{I} Regressionsanalyse abschätzen:

$$T_j = \rho^2(x_{-j}, T_j) X_j + [1 - \rho^2(x_{-j}, T_j)] E(X_j)$$

Dabei ist T_j die Zufallsgröße der wahren Testpunktwerte t_{jj} und X_j die Zufallsgröße der beobachteten Werte X_{jj} mit dem Erwartungswert $E(x_{-j})$, dem Populationsmittelwert. Die Größe $\rho^2(x_{-j}, T_j)$ ist die \hat{I} Reliabilität (s. u.) der Variablen X_j und kann abgeschätzt werden. Grundlegend für das weitere Konzept der klassischen Testtheorie ist der Begriff der parallelen Messung (f Äquivalenz von Messungen). Dieser Begriff, der rein formal über die Erwartungswerte und Varianzen definiert ist, bildet die Basis für das Reliabilitäts- und Validitätskonzept der klassischen Testtheorie. Die f Reliabilität einer Meßvariablen x_{-j} wird definiert als korrelativer Zusammenhang zwischen dem wahren und dem beobachteten Meßwert: $\rho(x_{-j}, T_j)$. Dabei bezeichnet ρ eine Produkt-Moment-Korrelation. Es gilt, daß alle parallelen Messungen jeweils paarweise einen konstanten und gleich großen Korrelationskoeffizienten haben und daß dieser identische Wert gerade mit dem Wert des Reliabilitätskoeffizienten übereinstimmt. Dieser Zusammenhang wird zur Schätzung der Reliabilität ausgenutzt, z. B. durch die Halbierungsmethode oder durch die Re-Test-Methode (f Spearman-Brown-Formel). Die \hat{I} Validität bezüglich eines Kriteriums Y wird ebenfalls als korrelativer Zusammenhang $\rho(x_{-j}, Y)$ definiert. Es gilt, daß alle parallelen Messungen die gleiche Korrelation mit dem Kriterium Y aufweisen, d. h. untereinander gleich valide sind. Stets ist ein hoch valider Test auch hoch reliabel.

Gegen diesen klassischen Modellansatz zur p. T. gibt es eine Reihe von Einwänden, die zur Weiterentwicklung des theoretischen Modells geführt haben. Wesentliche Kritikpunkte, z. B. nach FISCHER, sind:

1. Die problematische Voraussetzung einer metrischen Skala für die Beobachtungsdaten, die empirisch nicht entscheidbar ist. HORST hat eine Erweiterung speziell für dichotome Meßvariable entwickelt und begründet.

2. Die Voraussetzung des Parallelitätskonzeptes. Die Kritik liegt darin, daß dieser wesentliche Grundbegriff durch seine rein formale Definition keinen inhaltlichen Bezug hat. So ist es möglich, daß völlig unterschiedliche Meßvariable nach der formalen Definition parallel sind, während inhaltlich gleichartige Meßvariable mit unterschiedlicher \hat{I} Itemschwierigkeit nicht parallel sind. Zusätzlich