

AUS DER DIAGNOSTIK MAPPE: COMPUTERDIAGNOSTIK

EINIGE MÖGLICHKEITEN UND VIELE GRENZEN DER COMPUTERDIAGNOSTIK: (KUBINGER)

FUNKTION DES COMPUTERS IN DER DIAGNOSTIK:

- 1) Computer zur Vereinfachung der Auswertung
- 2) Computer zur Steuerung sogenannter peripherer Geräte
- 3) Verwendung des Computers zur Testvorgabe: Sie kann erfolgen auf drei Arten
 - 3.1 computerunterstützte Vorgabe von Tests
 - 3.2 Computervorgabe von PP-Tests
 - 3.3 Vorgabe originärer Computertests; man unterscheidet bei diesen 3 Arten:
 - 3.3.1 sogenannte Computer-Simulation
 - 3.3.2 mediumsgerechte Nutzung des Computers
 - 3.3.3 der „per Gelegenheit“ als Computer-Version erstellte Test

ad 1) Computer zur Vereinfachung der Auswertung:

- Computer erfüllt nur den **Ökonomie-Anspruch seitens des TL**;
- bietet **gewissen Komfort**;
- ABER: **höhere Kosten**;
- **Daten müssen selbst eingetippt werden** (Auswertungsfehlerquelle!) [Original-Protokollbögen einzuscannen ist zwar technisch möglich, aber noch nicht Standard, da sehr teuer]

ad 2) Computer zur Steuerung peripherer Geräte:

- für **Meßgeräte für neurophysiologische Messungen** (z.B. zur Bestimmung der Flimmerverschmelzungs-Frequenz);
 - für **Steuerung eines Tachiskops**
- ➔ **diagnostische Nützlichkeit ist aber noch nicht erwiesen**

ab 3.1 Computerunterstützte Vorgabe von Tests:

- expertensystemartige Hilfe für den TL bei der Testvorgabe

BEISPIELE:

- a) **adaptive Testvorgabe des AID** [ist aber noch nicht realisiert]: TL administriert den Test; Computer weist ihm dabei die nächsten Handlungsschritte an
- b) **DIASYS1** (computergestütztes normatives Diagnosesystem; gibt es, aber noch nicht publiziert; für Anwendungen im Bereich schulpsychologischer Beratungsstellen. Führt den Psychologen durch die verschiedenen Phasen des diagnostischen Prozesses, entlang allgemein akzeptierter und relevanter Wissensgrundlagen)

ad 3.2 Computervorgabe von PP-Tests:

- Übertragung von bewährten Papier-Bleistift-Tests auf den Bildschirm.
- **Kritik:**
 - ➔ es gibt **fast keine Studien zur Äquivalenzprüfung** [zwischen PP- und PC-Version]
 - ➔ vorliegende Studien zur Äquivalenzprüfung sprechen **eindeutig gegen die Verwendung der PP-Normen** (z.B. Kubinger entdeckte für RAVENS SPM im Durchschnitt 13 Punkte Unterschied zuungunsten der Computervorgabe; für Gießen-Test wiederum keine Normunterschiede; für IST-70 Unterschied zu gunsten der Computervorgabe)

Fazit: **Test müßte erst auf seine Äquivalenz in bezug auf die Normierung überprüft** werden, finanzieller Aufwand (Software-Anbieter wollen sich das nicht leisten)

ad 3.3.1 Computer - Simulation:

- erinnern an die frühen „Arbeitsproben“ oder an Computerspiele
- **Verhalten der TP soll in einer möglichst (berufs-)realistisch nachgeahmten komplexen Aufgabenstellung erfaßt** werden

BEISPIELE: -> **Mailbox 90** (dem „Postkorb“ des Assessment-Centers ähnlich)
 -> **AIRPORT**

Vorteil:

- ➔ **hohe Nützlichkeit**
- ➔ wegen **Erfassung des Realverhaltens** ist eher **prognostische Validität** zu erwarten als für herkömmliche Fragebögen (beziehen sich immer nur auf das Verbalverhalten!)

ad 3.2.2 mediumsgerechte Nutzung des Computers:

- darunter versteht man Vorgaben, für die der Computer nützlich, wenn nicht sogar notwendig ist
- **Notwendigkeit der PC-Verwendung kann bedingt sein**
 - > durch das Testmaterial,
 - > durch gleichzeitig erfaßten Item-Reaktionszeiten,
 - > durch erforderliche Rechenleistung

BEISPIELE:

a) adaptives Testen:

⇒ *branched testing:*

Itemauswahl pro TP erfolgt individuell in Abhängigkeit ihres Leistungsverhaltens bei vorangegangenen Items. Zur optimalen Auswahl müssen die Itemparameter (z.B. nach dem Rasch-Modell) bekannt sein (z.B. **BBT „Begriffs-Bildungs-Tests“**)

⇒ *tailored testing:*

Schätzung des gesuchten Fähigkeitsparameters ist nach jeder Itembearbeitung personenspezifisch notwendig (z.B. **Syllogismen**)

b) individuell verzweigend [heißt verwirrenderweise auch adaptives Testen...] z.B. Lerntests: **ADANA, Adaptiver Analogien Lerntest**

c) rechenbedingt mediumsgerecht wird der PC genutzt z.B. beim **Perseverationstest „Zeigeversuch“** (Rhythmusmessung mit Hilfe informationstheoretischer Maße)

d) materialbedingt mediumsgerecht wird der PC genutzt z.B. beim **Zeichen-Diskriminationstest** (2 zur Diskrimination vorgegebene Reize werden kurz dargeboten); **Tests zur Messung von Aufmerksamkeit und Konzentration**

ad 3.3.3 „Per Gelegenheit“ als Computer-Version erstellte Tests:

- haben nichts Computerspezifisches
- **hätten auch PP-Tests sein können**; sind bloß dem Trend der Zeit zuzuschreiben

BEISPIELE:

- ➔ **Topologischer Invarianztest**
- ➔ **Erkennen von META-Regeln**

GÜTEKRITERIEN VON COMPUTERTESTS:

- **Objektivität:**

- ➔ **Durchführungs- und Auswertungsobjektivität ist bei PC-Tests generell höher** als bei PP-Tests
- ➔ **Interpretationsobjektivität ist vollkommen gleich** (sofern es sich um normierte Tests handelt!) [Computer-“automatisierte“ Interpretationen sind zwar möglich, psychologisch aber NICHT zu vertreten -> Diagnostiker muß die im Laufe des diagnostischen Prozesses gesammelten Informationen gegeneinander abwägen, dabei kann ihm zwar der PC helfen, wissensbezogene Fehler zu vermeiden, ihn aber nicht ersetzen!]

- **Reliabilität:**

- ➔ **PC-Diagnostik ist überlegen**, wo mittels **(tailored)-adaptiven Testens** höhere Meßgenauigkeit erreicht werden kann (-> Hiermit ist auch ein Anspruch von **Ökonomie** erfüllt!)

- **Validität:**

- ➔ Validität von PC-Tests ist **nicht grundsätzlich besser**
- ➔ **Computer-Simulationen:** Wegen Erfassung des Realverhaltens ist **eher prognostische Validität zu erwarten als für herkömmliche Fragebögen** (beziehen sich immer nur auf das Verbalverhalten!)

- **Nützlichkeit:**

- ➔ **Computer-Simulationen:** Verhalten der TP soll in einer möglichst (berufs-)realistisch nachgeahmten komplexen Aufgabenstellung erfaßt werden; **hohe Nützlichkeit**
- ➔ **PC zur Steuerung peripherer Geräte** - diagnostische Nützlichkeit solcher Messungen ist nicht **erwiesen**
- ➔ **Computer-Version eines PP-Tests ist nicht nützlicher** als der PP

- **Normierung:**

- ➔ **Computertestungen sind Einzeltestungen**, die Datenerhebung kann nicht in Großgruppen erfolgen; es ist aber **so gut wie unmöglich ausreichend große UND repräsentative Stichproben zur Normierung zu erhalten** (Leute mögen nicht, es gibt rechtliche Probleme z.B. in Schulen; außerdem reduzierte Forschungsmittel, daher kein Geld für eine ordentliche Normierung)

- **Zumutbarkeit:**

→ PC-Tests sind hier **durchaus positiv** zu bewerten

- **Verfälschbarkeit:**

→ Frage stellt sich **nur bei Persönlichkeitsfragebögen**;

→ angeblich beantworten TP dem PC persönliche und intime Fragen offener als einem menschlichen TL (MENGHIN gelang der Nachweis, daß das nicht stimmt! 12 Tabu-Themen in einem herkömmlichen Fragebogen eingestreut -> **nur vereinzelte Unterschiede zwischen PP- und PC-Version**)

- **Fairness:**

→ Zwischen **Personen mit und ohne Erfahrung** im Umgang mit der Maus bestehen, wenn letztere entsprechend eingeschult werden **keine Unterschiede** (Beweis von HERGOVICH)

→ **Personen ohne Erfahrung könnten dennoch gehandikapt sein** (ist noch nicht hinreichend geklärt)

- **Skalierung:**

(= Feststellung der „Verrechnungsfairness“ der Testreaktionen zu Testwerten)

→ ist **nur bei Computer-Simulationen grundsätzlich problematischer** als bei PP-Tests (infolge der Item-Interdependenzen sind sie regelmäßig nicht ad-äquat)

→ **herkömmlichen Tests fehlt es meist am entsprechenden Nachweis**, obwohl das Methodeninventar (die Modelle der probabilistischen Testtheorie) vorhanden wären

VORTEILE DER COMPUTERDIAGNOSTIK:

⇒ Problem des multiple-choice Formats:

Bei **herkömmlichen PP-Tests** regelmäßig **hohe a-priori Wahrscheinlichkeit für Zufallstreffer** (hohe Ratewahrscheinlichkeit);

→ **Abhilfe durch PC-Tests**

BEISPIELE:

- beim **META**: **3 aus 6 gegebenen Antwortmöglichkeiten** müssen gefunden werden
- beim **Syllogismen**: **sequentielle Vorgabe** (TP muß für jede sequentiell gebotene Antwortmöglichkeit entscheiden, ob diese richtig oder falsch ist; entscheidet sie für „richtig“ [oder nach allen 4 Antwortmöglichkeiten], so erhält sie das nächste Item)

→ **Sequentielle Vorgabe ist (außer im Einzeltest) nur am PC möglich!**

- ⇒ Instruktion kann mit PC leicht nach dem Prinzip des „Programmierten Lernens“ gestaltet werden, sodaß die TP sie nachweislich versteht (bei PP-Tests fangen viele TP mit dem Test an, obwohl sie die Instruktion nicht verstanden haben, getrauen sich aber nicht das zuzugeben)
- ⇒ Es gibt bekanntlich verschiedene Wahrnehmungstypen; mit dem PC können die meist nur optisch vorgegebenen Instruktionen auditiv unterstützt werden (gibt es schon bei einem Test für die Deutsche Bundeswehr)
- ⇒ Bei herkömmlichen Persönlichkeitsfragebögen mit dichotomem Antwortformat („stimmt“ / „stimmt nicht“) wurde das Phänomen der Reaktanz festgestellt: TP reagieren auf die mit diesem Antwortformat verbundene Freiheitseingrenzung untypisch bzw. unwillkürlich und unter Umständen ihren wahren Eigenschaften zuwiderlaufend.

DAHER: **kontinuierliches Antwortformat verwenden (= Analogskala).**

Das ist **mit PC besonders ökonomisch** zu realisieren, egal ob Antworten nachträglich ohnehin dichotomiert werden oder nicht

GRENZEN DER COMPUTERDIAGNOSTIK:

- vor allem die Skalierung (= Feststellung der Verrechnungsfairness der Testreaktionen zu Testwerten):
 - meist fehlt es (wie auch den PP-Tests) am **entsprechenden Nachweis, obwohl die Methoden der Probabilistischen Testtheorie zur Verfügung stünden**
 - **Simulationsverfahren** bergen das Problem, daß **infolge der Item-Interdependenzen diese Modelle regelmäßig nicht adäquat** sind.