

## DREIDIMENSIONALER WÜRFELTEST (3DW)

### KURZCHARAKTERISTIK:

#### Testart:

- Niveau (**Power**) - Test zur
- Messung **des räumlichen Vorstellungsvermögens**;
- eindimensionaler (**rasch-homogener**) Fähigkeitstest;
- für **Einzel- und Gruppentestungen**

#### Altersbereich:

- **ab 13 Jahre** einsetzbar,
- keine Altersbegrenzung nach oben

#### Testformen und Testökonomie:

- \* **Langform:** vorgesehen für den **Standard-Einsatz**  
besteht aus einem **Warming-up Item** und **17 Testaufgaben**  
(Nr.2-18)  
ist ein **echter Power-Test** (ohne Limitierung der Bearbeitungszeit)  
→ kann fallweise sehr lang dauern; Testdauer nur grob einschätzbar

- \* **Kurzform:** da 3DW testtheoretisch in der probabilistischen Testtheorie fundiert ist, können **verschiedene Kurztestformen eingesetzt** werden.

**Abgestimmt auf** das jeweilige **Testziel** und die jeweilige **Testsituation**, können Kurzformen prinzipiell **auf 3 Arten hergestellt** werden:

- 1) durch Reduktion der Testaufgaben (**Testverkürzung**):  
Minimum: 8 Aufgaben
- 2) durch Limitierung der Bearbeitungszeit (**direkte Testabbruchmethode**):  
Minimum: 10 Minuten exklusive Instruktionszeit
- 3) durch Testabbruch, wenn die langsamste TP eine gewisse Aufgabenzahl bearbeitet hat (Minimum 8 Aufgaben)  
(**indirekte Testabbruchsmethode**)

Welche der 3 Methoden eingesetzt wird, entscheidet der TL

**Kurzformen haben:**

- \* etwas **geringere Meßgenauigkeit**
- \* dieselben positiven Testeigenschaften wie die Langform
- \* sind **praxisgerechter** Beitrag im Sinne einer ökonomischen, modernen Diagnostik

**Testinstruktion:**

für alle Testformen gleich, wird vom TL vorgelesen

**Testdauer:**

- ➔ 3DW wird je nach Motivation, Leistungsfähigkeit und individuellem Arbeitstempo der TP **unterschiedlich rasch bearbeitet**.
- ➔ Richtwerte bezüglich der **Langform**:
  - \* durchschnittliche Bearbeitungszeit = 22,44 Minuten (exklusive Instruktion)
  - \* Frauen: 22,85; Männer: 22,19 (vgl. Normstichprobe)
- ➔ Bei **Gruppentestung** ca. **35-40 Minuten** (werden einige TP nicht fertig -> direkte Testabbruchsmethode anwenden)
- ➔ Durch **Vorgabe der Kurzform**:
  - \* deutliche Reduktion der Bearbeitungszeit
  - \* Umgehung des Open-Ends der Power-Testung

**Testeinsatzbereiche:**

- Eignungsdiagnostik im Bereich der **Berufs-, Studien- und Bildungsberatung** und zur **Personalauslese**
- **Fähigkeitdiagnostik** in der psychologischen Praxis und im **klinischen Bereich** (z.B. Neurologie und Psychiatrie)
- für **Forschungszwecke**:
  - > Geschlechtsunterschiedsforschung
  - > Lateralisationsforschung
  - > ergopsychometrische Anwendungen

**Testtheoretische Analyse:**

- \* dichotomes logistisches Testmodell von **Rasch**
- \* lineares logistisches Test-Modell (**LLTM**) von Fischer

### Konstruktvalidität:

Nach neueren Untersuchungsergebnissen ist speziell **bei Würfelaufgaben** die dreidimensionale Aufgabenstellung noch **keine Garantie** dafür, daß in den TP die von der Testkonstruktion her intendierte

#### „Raumvorstellungsleistung“

(= vorstellungsmäßiges Manipulieren, Transformieren und/oder in Beziehung setzen von dreidimensionalen Objekten)

tatsächlich evoziert wird. Es müssen erst **ganz bestimmte Würfelaufgaben** (sogenannte **Flächenwürfelaufgaben**), die **mit einer Flächenstrategie bearbeitbar** sind, **ausgeschieden** werden, damit die 3D-komponente ein bestimmendes Element im Aufgabenbearbeitungsprozeß wird.

(vgl. 20 **Würfelaufgaben des IST** -> davon sind 10-13 Flächenwürfelaufgaben!)

Reliabilität: Reliabilitätskoeffizienten zwischen **0,82 - 0,91**

### Validität:

- im Rahmen von **LLTM-Analysen** Aussagen über die **Konstruktvalidität** (siehe oben)
- **Korrelationen mit anderen Tests, mit Schulnoten und Interessensgruppenvergleiche** ermöglichen differenzierte Beurteilung des 3DW bezüglich verschiedener Aspekte der Validität.

### Testauswertung:

pro Aufgabe erfolgt **Auswertung zweikategoriell** (richtig oder falsch).

**Rohwerte** werden ermittelt **durch Auszählen der richtigen Antworten**

→ können anhand von Tabellen umgewandelt werden in:

- \* **Rohwertschätzungen** (nur für Kurzformen)
- \* **Personenparameterschätzungen** (Fähigkeitsparameter des RM)
- \* **gruppenspezifische Prozentränge**
- \* **T-Werte**

### Eichstichprobe und Normierung:

insgesamt **5673 Personen** wurden in **Österreich** getestet

- **Schülernormen** (Geschlecht, Alter, Schultyp)
- nicht repräsentative Normen für **Studenten** (gesamt, getrennt nach Geschlechtern, Studenten technischer Universitäten vs. Studenten nicht-technischer Universitäten)
- **Vergleichstabellen für TP über 30 Jahre** (geringer Stichprobenumfang, werden daher nicht als Normen bezeichnet; getrennt nach Frauen und Männern; bei Männern außerdem Altersabstufung)

### **HISTORISCHE ENTWICKLUNG UND ERGEBNISSE:**

Faktorenanalytische Intelligenzforschung wurde begründet durch

**das Modell von SPEARMAN -> Generalfaktortheorie:**

Intelligenzmaße sind aufgrund eines, allen kognitiven Leistungen gemeinsamen Faktors (g-Faktor) und eines weiteren, testspezifischen s-Faktors erklärbar.

**Entwicklung faktorenanalytischer Techniken** brachte Zunahme des Interesses an der Erforschung verschiedener Intelligenzdimensionen mit sich. Beginn der **faktorenanalytischen Raumvorstellungsforschung** steht im Zusammenhang mit der **Konstruktion neuer Testmaterialien**, deren ursprüngliches Ziel die **Messung praktisch-technischer Fähigkeiten** war.

Um Vorerfahrung weitgehendst auszuschalten, wurde von allzu konkreten, mechanisch-technischen Aufgabenstellungen Abstand genommen -> für die den neuen Tests zugrunde liegende **Fähigkeitsdimension** kam die Bezeichnung „**Spatial Ability**“ (= Raumvorstellung) auf.

### **ZUM TESTTHEORETISCHEN KONZEPT DES 3DW: DARSTELLUNG UND ERGEBNISÜBERSICHT**

Im Rahmen der **Latent-Trait-Theorie** stehen **Testmodelle** zur Verfügung, die nicht nur über **wissenschaftstheoretisch hervorragende Eigenschaften** verfügen, sondern die auch **zur Bewältigung praktischer Meßprobleme geeignet** sind.

Im Vergleich mit den alten faktorenanalytischen Modellen sind **diese probabilistischen Testmodelle leistungsfähiger wegen**

- \* Stichprobenunabhängigkeit
- \* spezifisch objektiver Aussagen

Für die **Prüfung der Homogenität von Testmaterialien** und um **Aufschlüsse über die zugrundeliegenden Denkvorgänge** zu erhalten, **eignen sich** vor allem

- \* das dichotome logistische **Modell von Rasch** und
- \* das linear logistische Testmodell (**LLTM**) von Fischer

**GITTLER** machte **Möglichkeiten und Grenzen des speziellen Itemtyps der Würfelaufgaben** im Hinblick auf den Einsatz in der Raumvorstellungsdiagnostik deutlich.

Er führte sowohl psychologisch-inhaltliche als auch theoretisch-formale Aufgabenanalysen mittels probabilistischer Testmodelle durch

→ **3DW ist daher rasch-homogen**, d.h.

- er **mißt EINE** und zwar bei allen TP dieselbe **latente Fähigkeitsdimension** (damit ist Kritik an Raumvorstellungstests, sie würden bei Männern und Frauen aufgrund unterschiedlicher Lösungsstrategien nicht dieselbe Dimension messen, entschärft!)
- **Rohwert enthält ALLE relevanten Informationen zur Charakterisierung der TP bezüglich ihrer Raumvorstellungsfähigkeit** (d.h. wer Testwert 10 hat ist gleich leistungsstark)

### ENTWICKLUNG DES 3DW:

Im Rahmen einer lernpsychologischen Arbeit war es notwendig, ein **Testmaterial zu finden**, das verschiedenen Anforderungen gerecht werden sollte, eine davon war die **Homogenität der Testaufgaben nach dem Rasch-Modell**.

Dafür schienen die **Würfelaufgaben des IST** geeignet zu sein

→ es stellte sich aber heraus, daß sie **nicht rasch-homogen** waren (d.h. Subtest Würfel des IST erfaßt NICHT nur eine latente Dimension -> daher: Rohwerte können nicht als Indikatoren für die Raumvorstellungsfähigkeit interpretiert werden)

→ Nun ging Gittler der Frage nach, **warum die Aufgaben des IST nicht rasch-homogen** waren:

- wegen der **speed-Komponente**, die dem Rasch-Modell a priori widerspricht
- TP können im IST prinzipiell **zwei verschiedene Lösungsstrategien erfolgreich** einsetzen, von denen aber nur eine geeignet ist, die Raumvorstellungsdimension adäquat zu erfassen:

#### 1) Falsifikationsstrategie:

TP sollen einen Lösungswürfel aus den Antwortalternativen heraussuchen, indem sie die Identität des Vorgabewürfels mit dem Lösungswürfel erkennen

→ TP können dabei auch **ohne Identitätsprüfung zwischen Vorgabe- aller falschen Antwortalternativen die Testaufgaben und Lösungswürfel durch systematisches Ausscheiden richtig bearbeiten** und es bleibt dabei nur ein Würfel übrig, der halt dann die Lösung darstellt. **Falsifizieren des Antwortwürfels ist einfacher**, d.h. mit geringerem kognitivem Aufwand verbunden als die Feststellung der Identität zwischen Vorgabe- und Lösungswürfel.

#### 2) Verifikationsstrategie:

**Vorgabewürfel und Lösungswürfel werden verglichen** und Identität wird festgestellt.

#### Entwicklung neuer Testaufgaben:

\* Es muß dabei sichergestellt sein, daß die **Verifikationsstrategie die einzige Lösungsstrategie** darstellt.

- Neue Würfelaufgaben müssen **deutlich schwieriger** sein **als die IST-Aufgaben**, da eine **Power-Testadministration** vorgesehen war.

⇒ **GITTLER konstruierte 30 neue Würfelaufgaben**, die sich in Folgendem **vom IST unterscheiden**:

- 1) **zusätzliche Antwortmöglichkeit „kein Würfel richtig“**  
→ erfolgreiches falsifikationsstrategisches Vorgehen der TP wird dadurch unmöglich (Lösungswürfel muß mit Antwort 0 verglichen werden)
- 2) **Erhöhung der Aufgabenschwierigkeit** dadurch, daß es **pro Item 6 Antwortwürfel** gibt (im IST ist Lösungswürfel immer nur aus 5 gleichbleibenden Antwortalternativen herauszusuchen)
- 3) **Antwort H (Ich weiß die Lösung nicht)** -> dient bei geeigneter Instruktion und Motivation der **Minimierung von Rateeffekten**, wie sie Multiple-Choice-Format immer zu befürchten sind.
- 4) **Auf den 6 Seitenflächen jedes Würfels** sind **verschiedene, optisch gut unterscheidbare Muster**; es darf nicht vorkommen, daß auf einem Würfel zweimal dasselbe Muster vorkommt
- 5) Ist auf dem Lösungswürfel im Vergleich zum Vorgabewürfel ein neues Muster zu sehen, kann die TP nicht sicher sein, daß sich dieses auch auf dem Vorgabewürfel (halt verdeckt) befindet -> TP kann somit Identität der beiden Würfel nicht prüfen (Instruktion muß entsprechende Hinweise enthalten); **auf Antwortwürfeln A-f darf maximal EIN neues Muster aufscheinen.**

---

#### BEISPIEL:

ad 3DW - Items:

- \* links Vorgabewürfel,
- \* daneben 6 Würfel,
- \* daneben zwei Felder: „kein Würfel stimmt“ und „ich weiß es nicht“

\* Rohwert