

PERSÖNLICHKEITSDIAGNOSTIK

Diagnostische Verfahren **zur Erfassung von "Persönlichkeit"** (= von charakterlichen Eigenschaften) **beziehen sich kaum auf Theorien** (Unterschied zu Intelligenz-Tests!). Auch die **faktorenanalytisch begründeten Persönlichkeitsfragebogen** sind **selten zu einer anerkannten Theorie** konfirmatorisch konzipiert, sondern **per se konstruktbildend**.

Man unterscheidet:

- a) **Persönlichkeits-Fragebogenbatterien** (= mehrdimensionale Meßkonzepte, ähnlich den Intelligenztestbatterien). Sie werden unterteilt in:
 - 1) **faktorenanalytisch begründete Fragebogen** (Dimensionen werden empirisch festgelegt)
 - 2) **a-priori-"dimensionalisierte" Fragebogen** (ihre Dimensionen werden von vornherein, hypothetisch [also nicht empirisch!] festgelegt)
 - 3) **objektive Persönlichkeitstests** (hier geht es um die Registrierung von Verhaltensweisen, deren diagnostische Bedeutung der TP nicht unmittelbar einsichtig wird -> sind NICHT durchschaubar, d.h. keine Augenschein-Validität, daher Unverfälschbarkeit; gehen auf R.B.CATTELL zurück)
- b) **Spezielle Persönlichkeitstests** (messen spezielle, meist nicht-mehrdimensionale Persönlichkeitskonstrukte bzw. "neue" Persönlichkeitskonstrukte)
- c) **Projektive Verfahren**

Im Persönlichkeitsbereich ist die Psychologische Diagnostik durch pragmatische Überlegungen bestimmt

⇒ **HOFSTÄTTERS Überlegungen zur Differentiellen Psychologie:**

1. Wieviele (unabhängige) Dimensionen sind zur Darstellung der Unterschiede sämtlicher Individuen überhaupt notwendig?

- wird in jeder Dimension eine Abstufung in $k = 7$ Kategorien angenommen, so folgen daraus für eine Population von $P = 100$ Millionen $d = 10$ Dimensionen.

Erklärung dazu:

In einem Beschreibungssystem mit d Dimensionen und jeweils k Abstufungen, soll es für jedes Mitglied der Population P eine eigene Kategorie geben. Die Werte für d und k ergeben sich aus der Gleichung $P = k^d$. Bei gegebenen k [also $k = 7$] errechnet sich d aus der umgeformten Gleichung: $d = \log P / \log k$.

2. Wie komplex sind individuelle Beurteilungssysteme (d.h. wie differenziert beurteilt "der Mensch" Personen)?

➔ Dafür zieht Hofstätter **KELLYs Repertory-Prozedur** heran: Im ersten Schritt werden ziemlich viele Eigenschaften gesammelt; im zweiten Schritt werden alle "gridmäßig" (d.h. gittermäßig) zur Beurteilung einer Personengruppe verwendet (dabei wird für jeweils 3 bestimmte Personen der Gruppe diejenige Eigenschaft gesucht, die zwei der Personen haben, die dritte jedoch nicht). Aus seiner Erfahrung schließt Hofstätter, daß im wesentlichen dabei **3 Dimensionen** ausreichen:

- **kognitive Rekonstruktionen** einer Person von der Realität
- **emotionale Eindrücke** einer Person auf die Realität
- **verhaltensmäßiges Inventar** einer Person in der Realität

FRAGEBOGEN:

Antwortformate

1. **dichotomes** Antwortformat
2. **mehrkategoriell** (>2)
3. **Analogskala** (ohne Abstufungen; meist durchgehende Linien, auf denen die Lage der zutreffenden Antwort markiert werden soll)
4. **Q-Sort**

Kritik an dichotomen Antwortformaten

- Forced-choice führt zu **Mangel an Freiheiten bei der Abstufung**,
 - außerdem erscheinen **Fragebögen mit forced-choice-Format oft "unfreundlich"**.
 - Die DA von **KORNER** (1993) befaßte sich mit dem sozialpsychologischen Phänomen der **Reaktanz in Bezug auf dichotome Antwortvorgabe**. Diese konnte jedoch nicht nachgewiesen werden.
- ⇒ **Vorteil** ist der **hohe Informationsgehalt** bzw. die **hohe Informationsdichte** der Antworten **durch den Ausschluß von mittleren Antworten** ("weiß nicht").

Kritik an mehrkategorialen Antwortformaten

- Bei ungeraden Antwortformaten wird die **"Tendenz zur Mitte"** gefördert – de facto **Erhalt keiner Information**.
- Gibt es **zu viele Möglichkeiten** (z.B. sieben Stufen) kann das zu **Schwierigkeiten der Entscheidungsfindung** bei der TP kommen.

Kritik an Analogskalen

Durch Analogskalen gelingt eine (theoretisch unendlich) **feine Nuancierung der Antwort**.

Kritik an Q-Sort

- Die **Ursprünge der Q-Sort-Methode** liegen in der **Faktorenanalyse**.
- Aufgabe der TP ist es, die **Anzahl aller Items den Abstufungen zuzuordnen**.
- Meist wird die Freiheit in der Zuordnung jedoch eingeschränkt, indem **nur eine festgelegte Anzahl an Items einer Abstufung zugeordnet** werden können (z.B. ≥ 3 aber ≤ 7).
- Verwendung findet die **Q-Sort-Methode z.B. im BIT II** (Entscheidung über + bzw. – bei vier Wahlalternativen).

Am häufigsten wird in Persönlichkeitsfragebogen das **dichotome Antwortformat** verwendet.

Verfälschbarkeit (siehe auch S.94ff, Kap.4)

Definition:

→ *Das **Ausmaß**, in dem ein Test die **individuelle Kontrolle über Art und Inhalt der verlangten bzw. gelieferten Informationen** ermöglicht.*

→ In der **Leistungsdiagnostik** sind die **Tests nur in eine Richtung verfälschbar** – man kann sich nur schlechter machen, als man ist, nicht besser.

→ In der **Persönlichkeitsdiagnostik** kann man die **Antwort in beide Richtungen verfälschen**. Je **augenscheinlicher die Intention eines Fragebogens** ist, **desto eher** gilt es, dies zu **beachten**.

1. In einer Studie zur “Serviceorientierung” konnte gezeigt werden, daß **TP** mit dem Hinweis, gut abzuschneiden sämtliche Skalen verfälschten, um ein **möglichst gutes Bild von sich zu liefern** (KUBINGER).

2. In einer **Studie** mit 280 Propädeutikums-Teilnehmern wurde der **MBTI** (Meyer-Briggs Typen-Indikator, BENTS&BLANK, 1991) vorgegeben. Die KG hatte die Aufgabe, so ehrlich wie möglich zu antworten, die VG wurde informiert, daß die Antworten im Anschluß in der Gruppe diskutiert würden, negative Antworten also vermieden werden sollen. Auch hier konnte erwartungsgemäß ein Verfälschungseffekt festgestellt werden. **Im Bereich “feeling” (Empathie) war dieses Ergebnis besonders signifikant**. Offensichtlich nahmen die TP an, daß dieser Bereich für die Berufsausübung von immanenter Wichtigkeit sei.

3. PARFAHL (DA, 1998) fand **keinen Verfälschungseffekt** bei der Vorgabe des **EPP (Eysenck Personality Profiler)**. In Anbetracht der Tatsache, daß der Fragebogen 500 (!) Fragen umfaßt, könnte jedoch angenommen werden, daß die Instruktion im Lauf der Bearbeitung vergessen wurde.
4. **KECKEIS** (DA, 1999) fand den **Verfälschungseffekt auch bei Kindern**.
5. **RAAB-STEINER** (DA, 2000) fand, daß der **Verfälschungseffekt sogar unabhängig vom Antwortformat** gefunden werden kann.

Generell ist zu diesem Thema zu sagen, daß **Fragebögen in der TP Denkvorgänge provozieren**, die diese zu **Hypothesen über die Intention der Fragen** bringen. Diese muß jedoch nicht immer mit der Intention des Testautors übereinstimmen.

Faktorenanalytisch begründete versus a-priori-Fragebögen

Faktorenanalytisch begründet:

1. 16-PF (16 Personality Factors)
2. EPI (Eysenck Personality Inventory) bzw. EPP (Eysenck Personality Profiler)
3. GIESSEN-TEST
4. NEO-FFI (Neo-Five Factors Inventory)
5. FPI-R (Freiburger Persönlichkeits-Inventar, revidiert)

a-priori-Fragebögen:

1. MMPI (Minnesota Multiphasic Personality Inventory)
2. PRF (Personality Research Form)
3. TPF (Trierer Persönlichkeits Fragebogen)
4. MBTI (Myers-Briggs Typenindikator)

faktorenanalytisch begründete Fragebogen:

1) 16-PF (S.159)

⇒ **R.B. CATTELL** ließ die **18.000 Begriffe** (Adjektive für menschliches Verhalten) aus „Webster's International Dictionary“ (Studie Von **ALLPORT & ODBERT**) **nach subjektiver Zusammengehörigkeit ordnen**. Das Ergebnis wurde einer FA unterzogen, die daraus resultierenden **35 (bipolare) Faktoren auf 16 reduziert**. Eine **Kontrollgruppe** wurde gebeten, **sich mit den 35 Faktoren selbst zu beschreiben**, woraus **12 Faktoren** resultierten. Die **fehlenden vier Faktoren** wurden als **“questionnaire specific”** beschrieben (“Q1” bis “Q4”). (Sie sind für die Selbstbeschreibung spezifisch)

⇒ 1983 **deutsche Fassung des 16-PF** (SCHNEEWIND, SCHRÖDER & CATTELL)
 ➔ **seither heißen die 16 Faktoren:**

1. Sachorientierung vs. Kontaktorientierung
2. Soziale Anpassung vs. Selbstbehauptung
3. Besonnenheit vs. Begeisterungsfähigkeit
4. Flexibilität vs. Pflichtbewußtsein
5. Zurückhaltung vs. Selbstsicherheit
6. Robustheit vs. Sensibilität
7. Vertrauensbereitschaft vs. Skeptische Haltung
8. Pragmatismus Konkretes vs. Abstraktes Denken
9. Emotionale Störbarkeit vs. Emotionale Widerstandsfähigkeit
10. Pragmatismus vs. Unkonventionalität
11. Unbefangenheit vs. Überlegenheit
12. Selbstvertrauen vs. Besorgtheit
13. Sicherheitsinteresse vs. Veränderungsbereitschaft
 -> für Selbstbeurteilung spezifisch
14. Gruppengebundenheit vs. Eigenständigkeit
 -> für Selbstbeurteilung spezifisch
15. Spontaneität vs. Selbstkontrolle
 -> für Selbstbeurteilung spezifisch
16. Innere Ruhe vs. Innere Gespanntheit
 -> für Selbstbeurteilung spezifisch

⇒ **Insgesamt umfaßt der 16-PF 192 Items**

BEISPIELITEMS:

- Bei gleicher Arbeitszeit und gleichem Gehalt wäre es interessanter

a) Zimmermann oder Koch	0 Punkte
b) unsicher	1
c) Kellner in einem guten Restaurant zu sein	2
- "Spaten" verhält sich zu "graben" wie "Messer" zu

a) scharf	0
b) schneiden	1
c) Spitze	0
- Die Ziele, die ich mir in meinem Privatleben setze, erreiche ich fast immer.

a) Stimmt	2
b) Unsicher	1
c) stimmt nicht	0
- Ich vermeide es, Leute und deren Ansichten zu kritisieren.

a) ja	0
b) manchmal	1
c) nie	2

⇒ **In der deutschen Übersetzung gilt dies jedoch *nicht*.**

- Hier wurde **versucht, die selben 16 Faktoren zu finden, wie in der englischen Version (!?!).**
- Insofern **gelten für den deutschen 16-PF(-R) weder die Normen der englischen Version, noch das Adjektiv "faktorenanalytisch" begründet: Eigentlich** handelt es sich bei dieser Version des Fragebogens um *einen a-priori-Fragebogen!!!*
- Bei **wörtlicher Übersetzung der 4 Parallelfornen des Originals ergaben sich skalenweise zu geringe Interkorrelationen**

→ daher: ***Itemzusammenstellung der deutschen Version*** pro Skala **auf der Basis der Trennschärfen** (dabei wurde außerdem darauf geachtet, daß möglichst geringe Korrelationen der Skalen untereinander entstehen). Dann FA der endgültigen 16 x 12 Items (aufgeteilt in je 2 Halbskalen mit jeweils 6 Items).

Fazit: Weil bei der Erstellung der deutschen Version des 16-PF versucht wurde, exakt dieselben 16 Faktoren zu erfassen wie im Original, handelt es sich bei der deutschen Version des **16-PF eher um einen a-priori-Fragebogen als um einen faktorenanalytisch begründeten!**

Der **16-PF-R** (1998)

- beinhaltet **modifizierte Fragen und Faktorennamen**,
- im Gegensatz zum 16-PF sind die **Faktoren des 16-PF-R nicht bipolar**.
- Des weiteren muß angemerkt werden, daß der Scree-Plot der FA **nur mit sehr viel gutem Willen 16 Faktoren** zeigt.
- Wenn die **Simple-Structure-Kriterien angewendet** werden, resultieren **lediglich fünf (!) Faktoren** (siehe unten: Big Five bzw. Abb.44, S.160).

KRITIK:

- 1) 16-PF in deutscher Version ist ein Versuch, vorliegende Vielfalt von Forschungsinstrumentarien zum 16-PF zu vereinheitlichen -> **verbindliche Normen werden mitgeteilt;**
aber: 1. deutsche Ausgabe enthielt **keine umfassenden Reliabilitäts- und Validitätsuntersuchungen**; weitere Untersuchungen wären notwendig gewesen, wurden aber nicht durchgeführt (3. Version von 1994 enthielt diese angekündigten Untersuchungen bzw. deren Ergebnisse noch immer nicht!)
 ⇒ **zu wenig umfassende Reliabilitäts- und Validitätsuntersuchungen**
- 2) Art, wie in der deutschen Version die Items zusammengestellt wurden, ließ auf eine geglückte Konstruktvalidierung hoffen; ABER: trotz der Schlußfolgerungen, die die Autoren aus dem **Scree-Test** gezogen hatten, muß **an der Validierung gezweifelt werden**: Das „Eigenwert-Kriterium“ läßt erkennen, daß die Interpretation von mehr als 5 Faktoren wenig sinnvoll ist (daher gibt es ja auch nur 5 Sekundärfaktoren). Aufgrund des Scree-Tests hätte man sich für 5 (maximal 7) Faktoren entscheiden müssen
 ⇒ **zu viele Faktoren extrahiert bei der FA**
- 3) Autoren führten nach der FA eine **schiefwinkelige Rotation** durch -> hier muß stets **nach dem psychologischen Erklärungswert gefragt** werden. Im Streben nach wissenschaftlicher Ökonomie, im Sinn von Reduktion redundanter Datenstrukturen in informative, geht es in der ursprünglichen Zielsetzung der FA um den „wenigst-dimensionalen Raum“, der mit seinen RECHTWINKLIGEN Koordinaten alle Datenpunkte hinreichend genau abbilden kann.
Mit SCHIEFWINKELIGEN Koordinaten gelingt diese Abbildung genauer,
aber: man braucht dafür **mehr Dimensionen und die Interpretation korrelierender Faktoren fällt schwerer** (als die nicht korrelierender, wie bei einer orthogonalen Rotation) -> 16-PF gibt daher auch keine Anleitung, wie die Korrelationen zwischen den Faktoren bei der Interpretation des „Persönlichkeitsprofils“ zu interpretieren sind. Am Ende werden UNABHÄNGIGE Supra-Faktoren bestimmt...
 ⇒ **schiefwinkelige Rotation nach der FA: Faktoren sind nicht unabhängig voneinander und daher schwer interpretierbar**

- ⇒ ***Bestimmung von 5 Sekundärfaktoren, denen die 16 Primärfaktoren zugeordnet sind; das hätte man sich sparen können, wenn man nicht zu viele Faktoren extrahiert hätte!***
- ⇒ ***In der Praxis interpretiert Psychologe mehrere Testwerte immer wie unabhängige Größen*** (weil Mensch gar nicht in der Lage ist, die Testergebnisse korrelierender Faktoren entsprechend der genauen numerischen Größe der Korrelationen zueinander in Beziehung zu setzen und zu bewerten...)

Einschub: SCREE-TEST:

- = eine **von CATTELL vorgeschlagene Graphik.**
 - Alle **Eigenwerte der vollständigen Faktorenlösung werden der Größe nach aufgetragen.**
 - Wenn man nun beim kleinsten Wert beginnt, so zeigt sich ein **annähernd kontinuierlicher Verlauf, bis zu einem Knick** (der ist dort, wo der entsprechende Faktor nicht mehr bloß zufällig mehr Varianz (= Eigenwert) erklärt als der vorhergehende.
 - **Faktoren mit größeren Eigenwerten als der Knick-Faktor werden als inhaltlich bedeutsam interpretiert.** Wenn man nun den Eigenwertverlauf von unabhängigen, zu simulierenden Zufallsvariablen einzeichnet, dann sind alle Faktoren mit solchen Eigenwerten bedeutend, die oberhalb des Eigenwertverlauf der Zufallsvariablen liegen.
- ⇒ ***Beim 16-PF wurde diese Überprüfung nicht vorgenommen, denn sonst hätte es keine 16-Faktoren-Lösung gegeben!***

Einschub: Sedimentationshypothese (S.166)

- Die Sedimentationshypothese kommt **aus der Linguistik**.
- Sie besagt, daß es **in jeder Sprache für Objekte so viele differenzierte Begriffe** gibt, **wie es der Wichtigkeit des Objekts in der jeweiligen Gesellschaft entspricht**.
- Daraus folgt, daß **die mit dem Objekt verbundenen Begriffe und deren Konnotationen** in einer Gesellschaft **nicht die selben** sein können, wie in einer beliebigen anderen.
- Auf die Psychologie umgelegt **BEDEUTET DIES FÜR DIE ÜBERSETZUNG VON TESTS:**
 - ➔ **Begriffe**, die in der englischen (oder einer anderen) Sprache eindeutig etwas Bestimmtes bedeuten, **müssen im Deutschen nicht die selbe Bedeutung haben**.
 - ➔ Dies **gilt insbesondere für** die Entwicklung bzw. **Übersetzung von Persönlichkeitstest**.
 - ➔ Gerade für die Bezeichnung von Persönlichkeitseigenschaften sind **die mit den Begriffen verbundenen Konnotationen von immanenter Bedeutung**. Leichte **Abweichungen dieser Bedeutungsinhalte führen daher zu einem anderen Ergebnis**. Welchen Sinn hat es daher, Tests (und hier vor allem Persönlichkeitstests) in eine andere Sprache zu übersetzen?

EPI (Eysenck Personality Inventory), (S.161)

- im deutschsprachigen Raum **heute kaum mehr angewendet**
- **beruht auf GUILDFORD** (dessen Fragebogen gab es aber nie in deutscher Version!)
- EYSENCKs erster Fragebogen (**MMQ; Maudsley Medical Questionnaire**) wurde lediglich in seiner deutschen Form publiziert (1953). Die **Items** wurden **a priori ausgewählt** und **mit 18 Lügenitems aus dem MMPI ergänzt**. Zur Messung von „Neurotizismus“ konzipiert; 38 der dichotom zu beantwortenden Items beziehen sich auf körperliche Symptome (-> Neurotiker sind lt. Eysenck eher bereit, körperliche Schwierigkeiten anzugeben)
- **Nach FA** wurde der Fragebogen zum **MPI** und **maß** in dieser Form **auch Extraversion/Introversion**. (mit (24 Items auf dieser Skala)

- Eine **Bearbeitung diesen Fragebogens** führte zum **EPI** (Eysenck Personality Inventory), [je 24 Items auf den Skalen „Neurotizismus“ und „Extraversion / Introversion“ + Lügenskala mit 9 Items]

- **BEISPIELITEMS:**

- **Skala „Extraversion“:**

- Haben Sie gern Trubel und Betrieb um sich?
 - Sind Sie so schlagfertig, daß Sie einfach immer eine passende Antwort finden?

- **Skala „Neurotizismus“:**

- Geraten Sie oft in einen unruhigen Zustand, in dem Sie etwas tun wollen, aber nicht recht wissen, was?
 - Fühlen Sie sich ohne ersichtlichen Grund manchmal glücklich und manchmal traurig?

- dessen **Überarbeitung** 1998 zum **EPP** (Eysenck Personality Profiler) führte.

KRITIK:

- ⇒ besonders **geringe Breite der erfaßten Eigenschaften**
- ⇒ vergleichsweise **geringe Dimensionalität** des EPI
- ⇒ **fragwürdige Konzeption**
- ⇒ **theoretische Fundierung** vor allem des Neurotizismus-Konzepts ist **nicht voll befriedigend** gelungen
- ⇒ **wenig attraktiv für praktischen Einsatz**

Anmerkung: In der Diagnostik ist es **üblich** geworden, **nicht von Neurotizismus, sondern von Emotionalität zu sprechen** (*emotionale Labilität vs. emotionale Stabilität*)

GIEßEN-TEST (S.163)

- 1972 erschienen, von BECKMANN, BRÄHLER und RICHTER; 1991 revidiert; 1990 neu normiert
- **Ziel der Konstruktion:** Entwicklung eines objektiven und praktischen Instruments mit mittlerer Bandbreite
- **Unterschied zu allen anderen Persönlichkeitstests:** Einbeziehung von *sozialen Einstellungen und Reaktionen* (d.h. schwerpunktmäßige Akzentuierung psychosozialer Merkmale); daher: **für verschiedenartige gruppendiagnostische Fragestellungen geeignet**

- **Inhaltliches Konzept** zielt auf Items ab, die für die **Binnenstruktur und die psychosozialen Beziehungen der TP psychoanalytisch aufschlußreich** sind. Durch das **Selbstbild der TP** will man Wesentliches über ihre libidösen und aggressiven Impulse und deren Verarbeitung unter dem Einfluß von Ich und Über-Ich erfahren und die psychosozialen Tendenzen bzw. Abwehrformen der TP kennenlernen.
- **40 Items / 6 Skalen:**
 - (1) Soziale Resonanz
 - (2) Dominanz
 - (3) Kontrolle
 - (4) Grundstimmung
 - (5) Durchlässigkeit
 - (6) Soziale Potenz
- das **Antwortformat ist siebenkategoriiell**. ("Ich schätze, ich glaube, ich denke, ich habe den Eindruck, daß ich 3 2 1 0 1 2 3 daß ich nicht").
- **BEISPIELITEMS:**

(1) Ich schätze es gelingt mir eher schwer,	3 2 1 0 1 2 3	eher leicht, mich beliebt zu machen
(2) Ich habe den Eindruck, ich bin eher ungeduldig	3 2 1 0 1 2 3	eher geduldig
(3) Ich bin eher sehr wenig ordentlich	3 2 1 0 1 2 3	eher überordentlich
(4) Ich halte mich selten	3 2 1 0 1 2 3	oft für bedrückt
(5) Ich habe den Eindruck, ich gehe eher leicht	3 2 1 0 1 2 3	eher schwer aus mir heraus
(6) Ich glaube ich suche eher	3 2 1 0 1 2 3	ich meide eher Gesellschaft
- **Items der Skala 6** wurden nach Korrelationen mit den Variablen Schulbildung und Höhe des Einkommens ausgelesen. Skala 6 umfaßt **Aspekte des Erfolgs in der sozialen Umwelt**

KRITIK:

- ⇒ **nicht alle 40 Items** werden **in den lt. FA erstellten Skalen berücksichtigt**
- ⇒ **In FA resultierten nur 5 Faktoren; GT enthält aber 6 Skalen**
- ⇒ **Skala 6 ist nicht unabhängig von den anderen Skalen** (wegen des Konstruktionskriteriums: Korrelationen mit den Variablen Schulbildung und Höhe des Einkommens)
- ⇒ **genaue Angaben zur FA fehlen**
- ⇒ **Stichprobe ist sehr klein** (235 Patienten einer psychosomatischen Klinik)
- ⇒ Angesichts der stattgefundenen Skalenkonstruktion **wäre es überzeugender gewesen, den GT als a-priori-Fragebogen zu deklarieren.**
- ⇒ **Vorteile:**
 - **onomisches Verfahren**
 - **Möglichkeit des Einsatzes zur Fremdeinschätzung** -> Selbstbild und Fremdbild können diagnostisch verglichen werden
 - **Einsatz in der Paardiagnostik** (siehe Seite)

Einschub: Big Five

- Die **Faktorenanalyse** ist ein Verfahren, das **generell sehr stichprobenabhängig** ist.
- **Verschiedene Autoren** kamen **nach FA ihrer Daten** jedoch **immer wieder auf eine 5-Faktoren-Lösung** für ihre Daten (So wie auch das Urmaterial von CATTELL; siehe 4.3.1).
- Dies geschieht **unabhängig von den Fragebögen, dem Land oder der Sprache**, in welcher der Test verfaßt ist.

NEO-FFI (S.165)

- **COSTA & McCRAE** kamen ursprünglich zu einer **3-Cluster-Lösung**
 - ➔ Neurotizismus,
 - ➔ Extraversion,
 - ➔ Offenheit für Erfahrungen (open to experience),
- um dem "Anspruch" der Big Five zu genügen (!) wurden **noch zwei weitere Faktoren gefunden**
 - ➔ Agreeableness/ Verträglichkeit und
 - ➔ conscientiousness/ Gewissenhaftigkeit
 daher der etwas eigentümliche Name **NEO – Five Factors Inventory**.
- seit 1993 in deutscher Übersetzung, ABER: **noch immer keine Normen!**
- **Robustheit und Replizierbarkeit der 5 Faktoren-Struktur**

- Begriff „Offenheit für Erfahrungen“ sollte lieber „intellect“ oder „culture“ heißen

• BEISPIELITEMS:

- Ich bin nicht leicht beunruhigt
- Ich mag gern viele Leute um mich herum
- Ich mag meine Zeit nicht mit Tagträumereien verschwenden
- Ich versuche zu jedem, dem ich begegne freundlich zu sein
- Ich halte meine Sachen ordentlich und sauber

KRITIK:

- ⇒ **methodisch wurde sorgfältiger Aufwand** betrieben, um **Konstruktvalidität in bezug auf die Big Five zu belegen**
- ⇒ **keine Normierung für die deutsche Fassung**, daher (noch) **nicht zum routinemäßigen Einsatz in der Praxis geeignet**

Merke: **NORMEN, DIE AUF EINER GROßEN, ABER NICHT REPRÄSENTATIVEN STICHPROBE BERUHEN, SIND EIN HOLLER!**

FPI - R (Freiburger Persönlichkeits-Inventar)

- 1970 als FPI erschienene Fragebogenbatterie, erfreut sich **großer Beliebtheit**
- wird **zu den faktorenanalytischen Verfahren gezählt**; **ABER:** Autoren gehen von bestimmten psychologischen Konstrukten aus, die ihnen aufgrund ihrer Erfahrung und der Literatur als interessant erschienen. **FA dient zur Ordnung der Items** und als Entscheidungshilfe, wieviele Konstrukte näher ausgearbeitet werden sollten. **Scree-Test des Eigenwertverlaufs wurde zwar gemacht**, aber psychologische Interpretierbarkeit der Ladungsmuster und die verfügbare Anzahl markanter Items sind den Autoren praktisch wichtiger. Nach Festlegung der Anzahl der Faktoren -> Auswahl der Items, wurden aber auch wieder ausgetauscht, um schließlich für jedes Konstrukt Skalen in ausreichender Länge zu erhalten [**Items messen halt irgendwas...**]
- **inhaltliche Ausrichtung:** Anregungen dazu aus dem Arbeits- und Interessensbereich der Autoren:
 - psychophysiologisch orientierte Persönlichkeitsforschung
 - Evaluation von Psychotherapie und Rehabilitationsmaßnahmen vor allem bei psychovegetativen Syndromen und Herz-Kreislauf-Erkrankungen
 - Untersuchung über Belastungs-Beanspruchungsprozesse (Streß)
 - Bereich Aggressivität und Prosozialität

- **Aspekte der Lebenszufriedenheit und Beanspruchung** (Streß) scheinen auch in der Literatur gegenwärtig wichtige Rolle zu spielen [rein damit in den Test!] -> Autoren sehen in ihnen, ebenso wie in sozialer Orientierung (Prosozialität, Altruismus), Leistungsorientierung, Gesundheitssorgen eine sinnvolle und praxisbezogene Erweiterung des FPI
- **nicht berücksichtigt im Test sind stärker psychopathologische Skalen** (z.B. für Angst, Zwang, Psychotizismus), aber auch spezielle Verhaltensmerkmale wie **Kontrollüberzeugungen und Coping-Dispositionen**.

KRITIK:

- ⇒ In Anbetracht der stattgefundenen Skalenkonstruktion **wäre es besser gewesen, den FPI-R als a-priori-Fragebogen zu deklarieren**
- ⇒ dem FPI-R liegt überhaupt **kein allgemein bekanntes inhaltliches Konzept zugrunde**
- ⇒ **methodische Mängel** (siehe oben)

A-Priori-Fragebogen (S.167)

- Es gibt bis heute **sehr wenige originär deutschsprachige Fragebogen**. Einer der **bekanntesten und – unsinnigerweise häufig verwendeten – Fragebogen ist der MMPI**.
- **Tests für Kinder und Jugendliche innerhalb der Persönlichkeitsbatterien gibt es eigentlich nicht**; außer:
 - ➔ ***Problemfragebogen für 11-14Jährige***
 - ➔ ***Problemfragebogen für Jugendliche***

MMPI (Minnesota Multiphasic Personality Inventory; HATHAWAY & McKINLEY, 1943)

- Der MMPI ist **als klinischer Test gedacht**. Wenn dieser Test **überhaupt noch eingesetzt** werden darf (deutsch erstmals 1963), dann auch **ausschließlich im klinischen Bereich**. Eine **Vorgabe dieses Fragebogens in der Berufsberatung oder ähnlichem ist sinnlos**.
- Konstruktion **orientiert sich am Konzept der inhaltlichen Gültigkeit aus der Sicht klinischer Syndrome** (aus 1000 Items wurden diejenigen den entsprechenden Skalen zugeordnet, die sich als typisch herausstellten:
z.B. 76% der Depressiven stimmten einem Item zu / 22% der „Normalen“ taten das nicht -> Item wurde interpretiert als für Depressive entsprechend)

- Der MMPI besteht aus **566 Items**, von denen **350 mehr als einer Skala zugeordnet** werden (→ **hohe Iteminterkorrelationen**). Darüber hinaus ist der MMPI **mit 15 "Lügenitems" bereichert** worden.
- Die **Skalen** sind:
 1. Hypochondrie
 2. Depression,
 3. Hysterie,
 4. Psychopathie,
 5. Maskulinität/Femininität
 6. Paranoia
 7. Psychasthenie
 8. Schizophrenie
 9. Hypomanie
 10. Soziale Introversion/Extraversion

KRITIK:

⇒ ***aus einem Testknacker:***

- ➔ 566 Items / **viel zu umfangreich**; **Antwortkategorien**: „stimmt“, „stimmt nicht“, „zweifelhaft“
- ➔ Viele der Items enthalten **sehr seltsame Formulierungen** (z.B. „Ich glaube ich bin ein verdammter Mensch“; „Manchmal höre ich Stimmen.“ -> Dem wird wohl kaum eine TP zustimmen...), beschäftigen sich **mit Körpervorgängen** (z.B. „Einmal oder mehrmals im Monat habe ich Durchfall“, „Ich muß öfters als andere Wasser lassen.“ -> wer wird das schon z.B. bei einem Test anlässlich der Bewerbung um eine Stelle zugeben?); **stochern im Liebesleben der TP herum** (z.B. Ich träume viel von sexuellen Dingen“, „Ich fühle mich sehr stark von Personen meines Geschlechts angezogen.“)
- ➔ **Items sind so leicht durchschaubar**, daß die **Lügenskala auch nichts hilft**.
- ➔ **Männlichkeits- / Weiblichkeitsskala mißt mit seltsamen, realitätsfernen Maßstäben** (z.B. Ich gehe gern ins Theater -> „stimmt nicht“ ist die „falsche“ Antwort für einen Mann; gibt z.B. ein Mann zu, daß er nicht gern Soldat wäre, so wird er als feminin eingestuft); Festhalten am überkommenen Rollenklischee (Mann geht auf die Jagd, liebt laute Gesellschaften, Schlägereien -> Frau wird nur dann Führungsqualität bescheinigt, wenn die möglichst viele männliche Werte hat)

⇒ **Kubinger dazu:**

→ **Autoren sind sich sehr wohl der Verfälschbarkeit ihres Fragebogens bewußt** (vor der Interpretation der Testwerte, muß mittels von **4 „Validitäts“-Werten** geklärt werden, **ob der Test überhaupt ausgewertet werden darf:**

- Tendenz „einen normalen Eindruck zu machen“
- Tendenz, eine im Vergleich zu anderen ungewöhnliche Reaktion zu geben
- Ausmaß, in dem zu Statements nicht Stellung genommen wird (wie oft „weiß nicht“)
- Werte der Lügen-Skala

⇒ **zur Lügenskala:**

z.B. Ich würde beim Spiel lieber gewinnen als verlieren
 Manchmal verschiebe ich etwas auf morgen
 Ich sage nicht immer die Wahrheit

Um als aufrichtig zu gelten, sollte man **mindestens die Hälfte dieser Items bejahen** (zu geringer Lügenwert ist auch ungünstig -> TP gilt als ängstlich)

Fazit: MMPI ist ein **äußerst schlechter Test**, der aber **sehr oft eingesetzt** wurde (ist für Gruppenvorgabe geeignet -> hat aber nur **Normierungsstichprobe von N = 998 -> sehr klein!**)

PRF (Deutsche Personality Research Form; STUMPF et al 1985; S.172)

- Im Gegensatz zum MMPI ist die PRF nicht als klinisches Instrument, **sondern zur Personalauswahl entwickelt** worden.
 - Die **deutsche Fassung** ist eine gründliche **Überarbeitung der amerikanischen von JACKSON (1967) und bezieht sich ebenso auf die „Personologie“ von MURRAY (→ TAT):** Vorzug dieser Theorie = erfaßt das menschliche Verhalten weitgehend mit relativ einfachen, dem Allgemeinverständnis nahestehenden Konstrukten
- ⇒ Mensch ist ein aktiver Organismus, der nicht nur auf Druck der Situation reagiert, sondern bestimmte Situationen auch aufsucht und gestaltet. Erfäßt die Auswirkungsmöglichkeiten der jeweiligen Situation auf die Bedürfnisbelange und agiert dementsprechend.

- ⇒ Zwei zentrale Begriffe: „need“ (Bedürfnis) auf der Personenseite und „press“ auf der Situationsseite (ist der Zielzustand, den die jeweilige Situation erhoffen oder in Verlust zu geraten befürchten läßt) Das **Handeln** ist **durch „needs“ (≈Motive) bestimmt und von „presses“ (≈Situation) abhängig**. (d.h. needs und presses entsprechen einander thematisch -> ein need sucht sich sein press und umgekehrt)
- ⇒ **Murray überwindet damit die alternativen Motivationstheorien**: Dort Motivation zum Handeln entweder dispositionell (also traitmäßig) begründet oder durch situationsbezogenes Lernen.
- ⇒ An möglichen Motiven unterscheidet er zwischen **primären Motiven (Bedürfnisse) und sekundären, sogenannten psychogenen Motiven**

- **Skalen:**

- Leistungsstreben
- Geselligkeit
- Aggressivität
- Dominanzstreben
- Ausdauer
- Bedürfnis nach Beachtung
- Risikovermeidung
- Impulsivität
- Hilfsbereitschaft
- Ordnungsstreben
- Spielerische Grundhaltung
- Soziales Anerkennungsbedürfnis
- Anlehnungsbedürfnis
- Allgemeine Interessiertheit

- **ITEMBEISPIELE:**

- Ich ärgere mich über mich selbst, wenn ich etwas nicht gründlich gelernt habe
richtig / ? / falsch
- Ich versuche, so oft wie möglich in der Gesellschaft von Freunden zu sein.
richtig / ? / falsch
- Es macht mir nicht viel aus, wenn mich jemand in einer Diskussion aussticht
richtig / ? / falsch
- Ich versuche, andere unter meinen Einfluß zu bekommen, anstatt zuzulassen, daß sie mich kontrollieren
richtig / ? / falsch

- **Skala „Infrequenz“:**

- ➔ zur **Kontrolle der Antworttendenzen**;
- ➔ **10 Items** (z.B. Ich kann mit beiden Händen gleich gut und schnell schreiben.)

- endgültige **Itemauswahl aus einem Pool von 300 primär nach dem Konzept der inhaltlichen Gültigkeit** (mittels Experten-Ratings und Fremdeinschätzung von fiktiven Personen mit extremer Ausprägung in ausgewählten Konstrukten) Sekundär erfolgte die **Itemauswahl klassisch testtheoretisch**, und zwar nach innerer Konsistenz, Item-Schwierigkeit, konvergenter und diskriminanter Validität (insbesondere auf die soziale Erwünschtheit)
- **FA dienten nur der nachträglichen Absicherung der Skalen!**
- **Skala „Soziale Erwünschtheit“ (von JACKSON):** diskriminante Validierung der Items der einzelnen Konstrukte mit Hilfe des **„Differentiellen Reliabilitätsindex“:**
 = rechnerische Verminderung der **Trennschärfe eines Items minus Varianzanteil der durch die soziale Erwünschtheit erklärt wird.**
 (in deutscher Version ist diese Skala nicht enthalten!)
- **Übertragung des Tests auf deutsche Verhältnisse:**
 - ➔ es geht um **psychologische Äquivalenz der Übersetzung**, bzw. darum, ob die deutsche Version innerhalb ihres Kulturkreises ähnliche Eigenschaften aufweist wie die amerikanische Version in deren
 - ➔ 2 Äquivalenzprüfungen; 2 Validierungsstudien - dann Kürzung des Tests - 3 Validierungsstudien, Arbeit zur Verfälschbarkeit von PRF-Skalenwerten.
- Fazit:** Wegfall der Skalen „Selbsterniedrigung“, „Autonomiestreben“, Bedürfnis nach Abwechslung“, Sorgfalt“, „Mißtrauen und Beachtung von Sinnesqualitäten“
 ➔ **Überschneidungen und Redundanzen laut Autoren dadurch behoben**

KRITIK:

- ⇒ **Faktorenanalytischen Auswertungen auf Itemebene sind zu bemängeln**, weil **Faktorenanalyse auf dichotome Items angewendet**, immer zu **artifiziellen Faktoren** führt (andere Methoden gab es aber 1970 noch nicht)
- ⇒ Im Manual wird berichtet, daß die **14 Skalen nur 4 Faktoren begründen** (Leistungsorientiert-disziplinierte vs. impulsiv-spielerische Lebenseinstellung; Überlegenheitsstreben, Anlehnungsbedürfnis und Hilfsbereitschaft) -> **Bemühungen um möglichst unabhängige Konstrukte waren daher sinnlos.**

⇒

- ⇒ **MMTM-Ansatz (CAMPBELL & FISKE)** [zur *Konstruktvalidierung*] ist zwar **vorbildlich**, liefert **aber inhaltlich nicht voll befriedigende Ergebnisse**
- ⇒ Zur **Bestimmung der inhaltlichen Validität** ließ man **fiktive Personen fremdeinschätzen** -> das ist zwar eine gute Idee, **aber**: es wurden **nur wenige Konstrukte untersucht; nur 59 Beurteiler** (Offiziersanwärter)

TPF (Trierer Persönlichkeitsfragebogen; BECKER, 1989; S.175)

- **Theorie dahinter:**

Seelische Gesundheit (dafür gibt es 7 Indikatoren; ist die Fähigkeit zur Bewältigung externer und interner Anforderungen) und **Verhaltenskontrolle sind Superkonstrukte der Persönlichkeit** => dann Überprüfung auf empirischen Weg (Hauptkomponentenanalyse), ob diese A-priori-Einteilung beibehalten werden kann oder modifiziert werden muß

- Ausgangspunkt war das **Postulat bestimmter Indikatoren für seelische Gesundheit**.
- Nach Erhebung wurde eine FA durchgeführt und einzelne Skalen wurden revidiert. Insofern ist der **Test sowohl ein a-priori-, als auch ein faktorenanalytisch begründeter Fragebogen**. (also ein Kompromiß).

- **Skalen:**

- Verhaltenskontrolle
- Seelische Gesundheit
- Sinnerfülltheit vs. Depressivität
- Selbstvergessenheit vs. Selbstzentrierung
- Beschwerdefreiheit vs. Nervosität
- Expansivität
- Autonomie
- Selbstwertgefühl
- Liebesfähigkeit

- **Antwortformat 4-kategoriell, 118 Items.**

- **ITEMBEISPIELE:**

- Ich gehe großem Trubel und Aufregungen... aus dem Weg
immer / oft / manchmal / nie
- Wenn es etwas Wichtiges zu entscheiden gibt, weiß ich.... genau , was ich will
immer / oft / manchmal / nie
- Ich fühle mich ... einsam
immer / oft / manchmal / nie

- Ich denke ... über mich und mein Leben nach immer / oft /
manchmal / nie

KRITIK:

- ⇒ **es gibt** im Vergleich zu anderen Fragebogen **nichts Spezifisches zu kritisieren**
- ⇒ **Begrifflichkeit paßt Kubinger nicht** (Nähe der beiden ersten Skalen zu Extraversion/Introversion und Neurotizismus)

MBTI (MYERS-BRIGGS TYPE INDICATOR; 1962, 1991; S.176)

- ist **eindeutig ein a-priori Fragebogen**
- Das **theoretische Konzept** beruht auf der Typologie von JUNG:
 - ➔ Unterscheidung zwischen Funktions- und Reaktionstypen und vier Grundfunktionen denken, intuitieren, fühlen, empfinden.
 - ➔ Je nachdem, ob die **Libido mehr nach innen oder nach außen gekehrt** ist entscheidet man zwischen **Extraversion und Introversion**.
- Ursprünglich wurde **von BRIGGS eine eigene Typologie** entwickelt, die sich jedoch nicht sehr von der Jungs unterschied:
 - ➔ **Vier polarisierende "Präferenzen"**
 - Extraversion/Introversion;
 - Sinnliche/Intuitive Wahrnehmung,
 - Analytische/Gefühlsmäßige Beurteilung und
 - Beurteilung/Wahrnehmung**ergeben in Kombination 16 mögliche Typen.**
- Der Test besteht aus **90 Items**,
- **4 Skalen**:
 - Extraversion / Introversion
 - Sinnliche / Intuitive Wahrnehmung
 - Analytische / Gefühlsmäßige Beurteilung
 - Beurteilung / Wahrnehmung
- das **Antwortformat** ist **dichotom**.
- **BEISPIELITEMS:**
 - Wenn Sie mit mehreren Leuten zusammenkommen, was machen Sie gewöhnlich)
 - A beteiligen Sie sich am allgemeinen Gespräch ODER
 - B unterhalten Sie sich mit jemandem, den Sie gut kennen?
 - Kommen Sie gewöhnlich mit Menschen besser aus,
 - A die phantasievoll ODER

B die realistisch sind?

- Achten Sie eher

A auf die Gefühle anderer Menschen

ODER

B auf deren Rechte?

- Wenn Sie einem genauen Plan folgen

A sagte Ihnen das zu

ODER

B schränkt Sie das ein?

KRITIK:

- ⇒ den **Testautoren geht es** dem Konzept entsprechend **um eine Typisierung** der TP (d.h. um eine Identifizierung bzw. „Erkennung“), **nicht um normorientierte Diagnostik**; daher **keinerlei Normierung!**
- ⇒ **für den routinemäßigen Einsatz ist dieser Fragebogen daher ungeeignet!** (Beschreibung des Typischen erklärt ja nicht, wie oft dieses in der Bevölkerung realisiert ist)

Objektive Persönlichkeitstests

Definition:

- *Objektive Persönlichkeitstests sind Verfahren, die **unmittelbar das Verhalten eines Individuums in einer standardisierten Situation erfassen, ohne daß die Person sich in der Regel selbst beurteilen muß.***
- *Die Verfahren **sollen keine Augenscheinvalidität haben.** Das kann **durch die Aufgabenwahl oder bestimmte Auswertungsmethoden erreicht werden.***

- ⇒ Als Reaktion auf das **größte Problem in der Persönlichkeits-Diagnostik**, nämlich der **extrem hohen Augenscheinvalidität der Tests**, wird **seit R.B. CATTELL** daran gearbeitet, **wie man Tests so konstruieren kann, daß der interessierende Informationsgewinn resultiert, ohne daß die Person das Ergebnis in ihrem Sinn manipulieren kann.**

- ⇒ **Es gibt drei Gruppen objektiver Persönlichkeitstests** (S.177ff):

1. Sehr oft kommt es zum Einsatz von **Tests**, die **für die TP den Anschein eines Leistungstests** erwecken. In diesen Fällen ist eine **Verfälschung des Ergebnisses** durch die TP **beinahe ausgeschlossen**, da dieser nicht entdecken wird, was eigentlich die interessierende Variable ist, die mit dem Test gemessen wird. Wenn die Person z.B. die Größe zweier dargebotener

Flächen vergleichen soll und in Wirklichkeit die Zeit gemessen wird, welche die TP zum Fällern eines Entschlusses benötigt.

2. Die zweite Gruppe ist einem "typischen" Fragebogen ähnlich, gemessen werden jedoch **nicht die Antworten, sondern nur die Tendenzen**, Extreme – **also das formale Antwortverhalten**
3. Schließlich gibt es **Tests, die eigentlich ein Experiment darstellen**, indem der TP z.B. eine **stressreiche Situation, ein frustrierendes Erlebnis** etc. zugefügt wird und ihre **Reaktion darauf gemessen** wird (z.B. Nachlassen der Selbstsicherheit nach Frustration u.a.).

OA-TB 75 (OBJECTIVE ANALYTIC TESTS-BATTERY; dt. HÄCKER et al, 1975)

- Der OA-TB 75 besteht aus **50 (!) Skalen**, darunter
 - ➔ die **"Wechselbilder" von RUBIN** (Häufigkeit des Kippens läßt hoch im Faktor kortikale Wachheit; differenziert besonders gut zwischen Tpn, wenn die instruktionswidrige Aufforderung gegeben wird, sich gezielt um möglichst häufiges Wechseln zu bemühen. -> **kann möglicherweise die physiologische Aktivierbarkeit messen, es gibt aber keine Validierungsstudien dazu**)
 - ➔ oder **Fragen**, wie: „Wie lange brauchen Sie, um 5 Seiten in einem Abenteuerroman zu lesen - 2, 5, 10 oder 20 Minuten...“ und andere. (Testwert „Miß/Vertrauen in die eigene Leistungsfähigkeit [d.i. Summe der je nach Kategorie mit 1 bis 4 zu verrechnenden Antworten über alle Items] ist wahrscheinlich oft verfälscht; Testwert „Un-/Genauigkeit der Zeitschätzung (Annahme, wahre Zeit liegt in der Mitte -> **Tp kann danach beurteilt werden, wie sehr sie zu extremen Zeitschätzungen neigt**)
- Jedoch sind auch im Fall des OA-TB 75 **einzelne Untertests** enthalten, die **aus Persönlichkeitsfragebogen** stammen (z.B. 16-PF). Deren Items sind **zum Teil durchschaubar**. Andere Subtests haben im Stil von Umfragebögen freie Antwortformate (z.B. T 110: hier geht es um verschiedene Lebensziele)
- **Untertests der OA-TB 75 kann man in drei Gruppen einteilen:**
 - (1) Tests, die den **üblichen Leistungstests ähneln**: sind gekennzeichnet durch ihre Aufgabenstellung; sind sogenannte „performance-tests of personality“; Verfälschbarkeit ist praktisch ausgeschlossen

BEISPIEL: **Untertest T 45:**

TP muß unter Zeitdruck die Längen von 42 Linienpaaren vergleichen; „Schnelligkeit und Genauigkeit des Urteils“ soll damit erfaßt werden -> das soll Indikator für **„Selbstbehauptung“** und **„Überschwenglichkeit“** sein; ABER: es gibt **keine Validierungsstudien dazu!**

- (2) Tests, die **herkömmlichen Fragebögen ähneln**;
Auswertung bezieht sich aber auf das formale und nicht auf das inhaltliche Antwortverhalten der Tp; dadurch soll Durchschaubarkeit des Meßprinzips reduziert bzw. verhindert werden

BEISPIEL: **Untertest T 19(1):**

Grad der Übereinstimmung zu 20 Statements wird erfragt (5 kategoriell von „stimme völlig“ bis „stimme überhaupt nicht überein“;) Statements sind aber jeweils ergänzt um eine Begründung für bzw. gegen die Zustimmung (z.B. Wenn jemand stiehlt, sollte er strengt bestraft werden -> Stimme völlig überein, denn wenn einer stiehlt und dafür nicht bestraft wird, wird er es immer wieder tun). Testkennwert mißt **die „Tendenz zur Übereinstimmung mit Urteilen“**

- (3) Tests, die **nicht in das Schema Leistungstests vs. Persönlichkeitsfragebögen passen:**

BEISPIEL: **Subtest T 17:**

analog zu den üblichen Kodiertests; TP soll 4 Buchstaben 4 Symbole zuordnen; Liste wird in 5 Teilzeiten vorgegeben; nach jeder Liste muß Tp eine Prognose darüber abgeben, wieviele Symbole sie in der nächsten Teilzeit kodieren wird; Testkennwert = Kodiermenge; durchschnittliche Differenz zwischen erzielter Leistungsmenge und der Prognose (**„Anspruchsniveau“**) und der durchschnittlichen Abweichung zwischen prognostizierter und tatsächlich erzielter Leistungsmenge (**„Zieldiskrepanz“**); ABER: **Kriteriumsvalidität ist ungeklärt.**

KRITIK:

- ⇒ es gibt für die OA-TB 75 **noch immer keine Normierung**
- ⇒ fast allen Untertests **mangelt es an Untersuchungen zur Skalierung**
- ⇒ **nicht ausreichende Validierung**
- ⇒ Testbatterie als Ganzes ist **wegen ihrer Länge unzumutbar**; ist daher **für die Praxis ungeeignet!**
- ⇒ ist aber eine **gute Quelle für Ideen zu objektiven Testkonzepten**

Merke: Generell ist anzumerken, daß der Versuch, einen **Test mit niedriger Augenscheinvalidität zu entwickeln, für den TL angenehm** ist, da er darauf vertrauen kann, die **tatsächlichen Persönlichkeitsmerkmale der TP zu erfahren**
→ die **ethische Unverträglichkeit** sei hier nur angedeutet

Weitere Ansätze

1. Farbe-Wort-Interferenz-Test (FWIT, STROOP):

- Die TP hat die Aufgabe, eine **Liste von farbig gedruckten Worten „vorzulesen“** und zwar, indem sie die **Farben benennt, in denen die**

Worte geschrieben sind und nicht die Farben, für welche die Worte stehen (z.B. "Blau" in roter Farbe geschrieben – die Person muß "blau" sagen).

- **Testkennwert** = Zeit, die die Tp zum Durcharbeiten der Liste braucht
- **Kann sie das nicht**, soll das ein **Anzeichen für ein organisches Psychosyndrom, Psychosen, Perseverations-Tendenz, Motivation** etc. sein.

KRITIK:

⇒ **dürftige Validierungsstudien**, daher ist Test mit Skepsis zu betrachten

2. Matching Familiar Figures (KAGAN):

- Die TP hat die Aufgabe, ein **Bild, das mit einer Vorlage identisch ist, so schnell wie möglich zu finden**. 12 Items (ähnlich wie der Subtest „Beobachtung“ des WIT)
- Gemessen wird die Zeit, bis die Entscheidung getroffen wird, und die Fehlerzahl, woraus **Rückschlüsse auf "Impulsivität versus Reflexivität"** gezogen werden. (Ist TP in Problemsituationen eher langsam und fehlerarm oder schnell und fehlerreich.)

KRITIK:

⇒ bei Kindern erprobt, aber es **fehlt ein überzeugender Validierungsbefund**, vor allem was die Generalisierbarkeit des mit dem Test gemessenen „kognitiven Stils“

3. „Repression vs. sensitization“:

- vgl. **Wahrnehmungsexperimente von BRUNER & POSTMAN:**
Es gibt Personen, die bei tachiskopischer Darbietung von emotional- bzw. angstbesetzten Wörtern gegenüber neutralen Wörtern verlängerte Erkennungszeiten (**Repressoren**) oder verkürzte Erkennungszeiten (**Sensitizer**) haben
-> **entspricht den FREUDschen Abwehrmechanismen**
- **ABER:** **Umsetzung eines solchen Tests auf PC ist zur Zeit noch nicht möglich** (zu großer technischer Aufwand); anders geht das aber nicht, wenn man einen routinemäßig einsetzbaren Test haben will

4. „Rigidität im Denken“:

- **ausgehend von LUCHINS „Umfüllaufgaben“** sollte auch die Rigidität im Denken objektiv testbar sein
- bei diesen Umfüllaufgaben zeigte sich, daß Menschen mehr oder weniger dazu neigen, **einmal gefundene Lösungswege beizubehalten, selbst wenn es alternative, weniger komplizierte Lösungswege gäbe**
- Für so einen Test im Rahmen der Persönlichkeitsdiagnostik **bräuchte man aber mehr als nur Umfüllaufgaben**
- **Es gibt bereits einen „Rigiditätstest“ (von KUBINGER)**, aber nur wenige empirische Erfahrungen dazu.

5. Gestaltwahrnehmung bzw. Feld- (Un-)Abhängigkeit:

- Die TP soll **Eckpunkte eines einfachen Haussymbols mit der Maus markieren.**
- Dieses wird **in einen immer komplexer werdenden Hintergrund verschiedener Linien eingebaut.**