

# Wissenschaftliche und praktische Gütekriterien

Dr. W. Strasser

August 1989

## INHALTSVERZEICHNIS

1	Ist eine Personalauswahl mit Test überhaupt möglich?
2	Verantwortung in der Eignungsdiagnostik
3	Grundlagen und Konstruktionsideen des vorliegenden Tests
4	Kriterien für die Beurteilung von Auswahlssystemen
5	Kriterium Objektivität
6	Kriterium Reliabilität
7	Validität
8	Vergleich: Validitäten von anderen Meß-Methoden
9	Validitätsstudien
10	Validitätsstudien im europäischen Vergleich
11	Akzeptanz
12	Anhang I        Einige Item-Beispiele aus anderen Tests
13	Anhang II        Einige Daten zum Test
14	Anhang III        Mittelwerte / Standardabweichungen für die einzelnen Länder
15	Anhang IV        Verteilung der Antworten auf die einzelnen Alternativen jedes Items
16	Literatur- und Abkürzungsverzeichnis

## *Abstract*

Aus mehreren Faktorenanalysen über „außerfachliche Kompetenzfaktoren“ wie Leistungsmotivation, Stabilität, Belastbarkeit, Extraversion wurde ein Fragebogen mit 157 Items zusammengestellt. Aus den ursprünglichen Analysen wurden insgesamt 16 Faktoren ausgewählt, die mit verschiedenen Aspekten des Berufserfolgs korrelierten.

Neben Validierungen mit unterschiedlichen Außenkriterien in Deutschland wurden auch Adaptationen in europäischen Sprachen mit dem Vorgesetzten-Urteil validiert. In allen Fällen ergaben sich substantielle Validitäten, die eine Beziehung des Tests für Auswahlentscheidungen rechtfertigen.

## 1. IST EINE PERSONALAUSWAHL MIT TESTS ÜBERHAUPT MÖGLICH?

Angesichts der Komplexität der Zusammenhänge zwischen psychologischen Eigenschaften und Leistungs-/Erfolgskriterien wird oft die Meinung geäußert, dass es nicht möglich sei, durch psychologische Tests Erfolg vorherzusagen.

Eine solche Einstellung ist einerseits falsch, andererseits auch bedauerlich, insbesondere dann, wenn sie von Wissenschaftlern oder Journalisten geäußert wird. Es ist ganz klar, dass Fortschritt in allen Gebieten, sei es in der Technik, Wirtschaft, Medizin oder Psychologie bislang einzig und allein über den Weg immer exakterer, quantitativer Erfassung von Einfluss-Parametern möglich wurde. Es wäre absurd, dies abzustreiten. Sicherlich sind in einigen Gebieten der Potentialanalyse noch keine besonders ermutigenden Ergebnisse vorhanden, aber auf der anderen Seite gibt es in einigen Gebieten durchaus zufriedenstellende Resultate, die gerade durch ihre exakte quantitative Erfassung immer weitere Verbesserungsmöglichkeiten erschließen.

Dies gilt um so mehr, als die traditionellen Methoden der Auswahl mit vielen Fehlerquellen belastet sind und, wissenschaftlich auf ihre Gütekriterien überprüft, sehr schlecht abschneiden. Will man nicht mit zweierlei Maß messen, dann dürften schriftliche Bewerbungsunterlagen und Interviews alleine aufgrund ihrer schlechten Voraussagequalitäten niemals als alleinige Einstellungskriterien verwendet werden. Aufgrund der hohen Verantwortung bei der Eignungsdiagnostik sollte aber kein Aufwand gescheut werden, um zu einem insgesamt befriedigenden Ergebnis zu kommen.

## 2. VERANTWORTUNG IN DER EIGNUNGSDIAGNOSTIK

### 2.1. Verantwortung für den Bewerber

- o Den Bewerber richtig beraten/Sicherheit geben,  
keine falschen Hoffnungen machen.  
Nicht versuchen, unqualifizierte Bewerber zu "retten".  
Dem Bewerber ist nicht gedient, wenn er Misserfolge erleidet, unter  
großem Stress wegen Überforderung steht, einen "dunklen Punkt" in  
seiner beruflichen Laufbahn hat.
  
- o Jedem Bewerber die optimale Chance geben, seine wirklichen Talente  
darzustellen.  
Ohne unbeeinflussbare Antipathie-Effekte,  
ohne Prüfungs-Stress.
  
- o Jeden Bewerber seriös behandeln.  
Keine unfairen Methoden.  
Keine Methoden, die den Beurteilungskriterien nicht genügen.

### 2.2 Verantwortung für die ökonomisch richtige Entscheidung dem Unternehmen gegenüber

Einstellungen neuer Mitarbeiter sind sehr teure Investitionen, die langfristig bis in Millionenbeträge gehen. Es sollte alles ökonomisch und ethisch vertretbare unternommen werden, um diese Investitionen so weit als möglich abzusichern.

### 2.3. Was kosten schlecht arbeitende Kräfte?

- Sie machen schlechte, teure Entscheidungen
- Sie hemmen Wachstum und Innovation
- Sie demoralisieren Kollegen und Mitarbeiter, so dass deren Wirkungsgrad sinkt und Unzufriedenheit entsteht,
- damit sinkt auch die Produktivität
- Es entstehen hohe Kosten durch Fluktuation
- Zeit und Energie werden auf Abwehr-, Vertuschungs- und Rache Maßnahmen verschwendet
- Das Top-Management muss teure Zeit für diese Dinge opfern.

## 2.4. Unwissenschaftliche Auswahlmethoden

Wenn keine wissenschaftlich überprüften Verfahren zur Erfassung des Potentials vorliegen, dann werden häufig als Ersatz ungeeignete Methoden benutzt:

1. Graphologie
2. Astrologie
3. Überinterpretation von Bewerbungsschreiben
4. Stress-Interviews
5. "Einfach-Indikatoren" ("Krummer" Lebenslauf als alleiniges Kriterium)
6. "Selbstgestrickte" Assessments
7. Tests aus anderen Bereichen werden wahllos übernommen (MMPI, Schulintelligenz-Test, FPI, usw.)

### 3 GRUNDLAGEN UND KONSTRUKTIONSIDEEN DES VORLIEGENDEN TESTS

Die grundlegende Idee des Verfahrens war, einen Test zur Erfassung von Vorbedingungen des Berufserfolgs auf einer breiten Basis zu schaffen. Dabei wurde von der Annahme ausgegangen, dass es nicht einzelne "Wunderfaktoren" gibt, die die Grundlage des Erfolgs sind, sondern mehrere Einstellungsfaktoren, wie das Selbstvertrauen, die Leistungsmotivation, Ängste im Leistungsbereich sowie Extraversion und andere. Von vornherein wurde nach berufsbezogenen Grundlagen gesucht, nicht nach allgemeinen Persönlichkeitsstrukturen. Es kristallisiert sich offenbar immer mehr heraus, dass allgemeine Persönlichkeitseigenschaften relativ wenig mit dem Berufserfolg zu tun haben; nicht aber die hier erfassten "außerfachlichen Kompetenzfaktoren". Der Unterschied zwischen beiden kommt schon in der Gestaltung der Items zum Ausdruck. (Als Beispiel für berufs-unspezifische Persönlichkeitsfragebögen sind im Anhang einige Items des 16PF-Fragebogens wiedergegeben.) Das Ergebnis mehrerer intensiver Faktor- und Itemanalysen in verschiedenen Zusammensetzungen ist der vorliegende Fragebogen.

## 4. KRITERIEN FÜR DIE BEURTEILUNG VON AUSWAHLSYSTEMEN

### 4.1. Klassische testtheoretische Kriterien

1.1. Objektivität

1.2. Reliabilität

1.3. Validität

### 4.2. Kriterien der praktischen Verwendbarkeit

2.1. Akzeptanz der Methodik

2.2. Datenschutz

2.3. Test als Förderungs- und Entwicklungs-Instrument

5. Kriterium Objektivität

Objektivität bedeutet:

Durchführung des Tests

unter vergleichbaren Bedingungen

Auswertung des Tests/der Informationen

nach festen Methoden und Standards, so daß zufällige Varianten oder Fehler ausgeschlossen werden

Interpretation der Daten

nach festen Regeln, so dass keine Fehlerquellen entstehen können

Diese Bedingungen sind im Gegensatz zu den meisten anderen Methoden (Interview/Assessment, etc.) voll erfüllt.

6 Kriterium Reliabilität  
(Verlässlichkeit)

Was ist Reliabilität?

Damit wird die Tatsache ausgedrückt, dass eine Messung zuverlässig ist, d.h. zu unterschiedlichen Zeitpunkten bei den gleichen Personen vergleichbare Messwerte ergibt.

Die Reliabilität setzt sich zusammen aus:



### Welche Koeffizienten sind zu erwarten?

Wenn die Reliabilität gegen 1,00 geht (1,00 wird wohl in der Praxis nie erreicht), dann hat man sehr zuverlässige Messdaten. Allerdings bedeutet das auch, daß ein Faktor dann extrem stabil, also nicht veränderbar ist. Bei extrem stabilen Faktoren lohnt sich in der Regel kein Training oder sonstige Maßnahmen, da keine Beeinflussbarkeit vorhanden ist.

Bei Kompetenzfaktoren, die für den Berufserfolg eine Rolle spielen und gleichzeitig auch noch durch Training veränderbar sein sollen, sind Reliabilitäten von 0,60 bis 0,80 gute Werte

## Ergebnisse

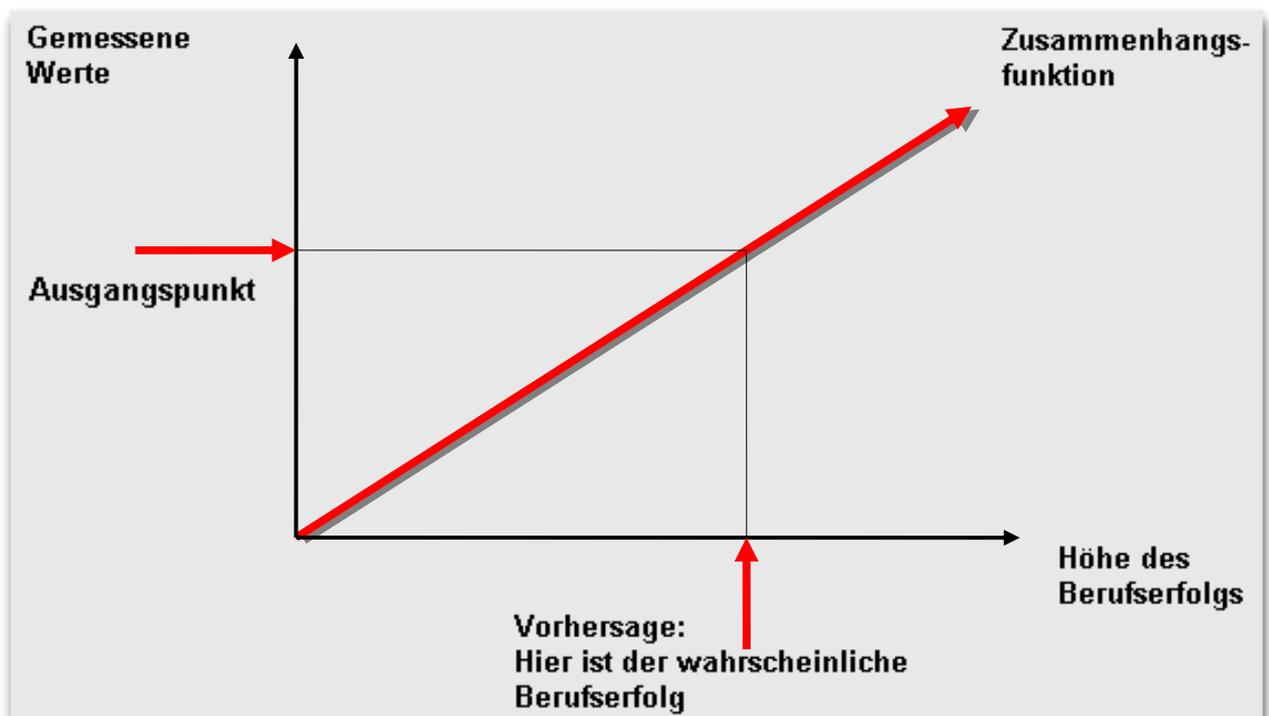
In verschiedenen Untersuchungen  
(z.B. SEDLMAYR, 1969, STRASSER, 1975)  
wurden Reliabilitätskoeffizienten zwischen 0,60 und 0,90  
mit einem Median-Wert von  
0,83 für einzelne Faktoren gefunden.

Die gefundenen Werte garantieren eine sehr  
gute bis ausreichende Reliabilität der  
Testwerte für das DNLA-Verfahren.

7. Validität  
(Gültigkeit)

7.1. Was ist Validität?

Mit Hilfe von Testwerten soll der wahrscheinliche Berufserfolg vorhergesagt werden. Grafisch ausgedrückt:



Die Validität wird wie die Reliabilität mit einem Korrelationskoeffizienten ausgedrückt, der zwischen  $-1,00$  und  $+1,00$  variiert. Keine Korrelation (kein Zusammenhang) wird mit dem Wert  $0,00$  ausgedrückt; ein vollständiger Zusammenhang, also hundertprozentige Vorhersage, bedeutet einen Koeffizienten von  $1,00$ .

Validität ist eigentlich das zentrale Kriterium für die Güte eines Tests. Wenn kein signifikanter Zusammenhang zwischen Testwert und Erfolgskriterium besteht, dann ist es besser, keinen Test zu verwenden. In einem solchen Fall ist auch das subjektivste Urteil eines Interviewers vermutlich besser als die (scheinbar) objektiven Testwerte.

## 7.2. Verschiedene Aspekte der Validität

### "Theoretische Validität" (Konstrukt-Validität) der gemessenen Merkmale

Darunter soll erstens die Tatsache verstanden werden, dass die im Test gemessenen Merkmale mit den von Fachleuten geäußerten Berufsvoraussetzungen/-anforderungen übereinstimmt.

Zweitens kann man systematische Beobachtungen über das Verhalten von Personen mit hohen/niedrigen Ausprägungen der jeweiligen Merkmale anstellen. Dies ergibt ebenfalls deutliche Hinweise darauf, dass die gemessenen Merkmale offenbar einen, starken Zusammenhang zum Erfolg-/Misserfolg haben.

### Gemessene Validität

Darunter versteht man quantitative Ergebnisse von Validierungsstudien.

### Praktische Validität

Ausgedrückt durch die Erfahrungen im praktischen Einsatz des Verfahrens.

### 7.3. Validität beginnt bei der- Konstruktion eines Tests

Tests können auf sehr unterschiedliche Arten und Weisen konstruiert und validiert werden.

#### Methode 1:"Voll-empiristische Methode"

Ein sehr großer Pool von unausgelesenen Items wird einer Stichprobe von Erfolgreichen (E) und Nicht-Erfolgreichen (NE) gegeben.

Die Items, bei denen sich signifikante Unterschiede ergeben, werden zu einem Test zusammengefasst.

Dabei können u.U. so "sinnvolle" Kriterien herauskommen, dass der Konsum von mehr als 5 Tassen am Tag ein Erfolgskriterium ist. Dies liegt daran, dass bei dieser Methode nicht unterschieden wird zwischen Erfolgs-Ursachen und Konsequenzen. Einige biografische Informationsbögen werden nach diesem Muster aufgebaut.

#### Methode 2: Faktoren-Analytische Methode

Ein sehr großer Itempool wird durch Faktorenanalysen strukturiert, reduziert und auf Unterschiede zwischen Erfolgreichen und Nicht-Erfolgreichen analysiert.

Problem 1: Diese Methodik gibt insgesamt nur gute Ergebnisse, wenn der Itempool alle Faktoren abdeckt, die für eine Position wichtig sind. Wenn das nicht der Fall ist, dann entstehen mehr oder weniger zufällige Faktoren der Persönlichkeit, die insgesamt schlecht in psychologische Gesamtzusammenhänge einzuordnen sind.

Problem 2: Wenn ein solcher Test zur Bearbeitung einer bestimmten Fragestellung konstruiert wurde (z.B. Neurosen-Diagnose) und dann auf andere Gebiete (z.B. Verkaufserfolg) übertragen wird, entstehen völlig unbrauchbare Ergebnisse

### Methode 3: Test-Konstruktion nach Systemanalyse

Die Zusammenhänge zwischen Leistungen/Leistungsergebnissen und psychologischen Variablen werden gründlich analysiert (Interviews, Beobachtungen, etc.) und die wichtigsten Leistungsvoraussetzungen in Faktoren erfasst. Anschließend werden Tests konstruiert, die diese Kriterien erfassen. Entwicklung der Tests nach den gängigen test-theoretischen Grundlagen.

Mit dieser Methode ist die Validität sozusagen "hineinkonstruiert", nicht erst nachträglich "herausanalysiert", wie bei Methode 1 und 2. Sie bringt deshalb die besten Ergebnisse.

#### 7.4. Konstrukt-Validität

Berufserfolg ist eine Größe, die vermutlich aus vielen unterschiedlichen Faktoren zusammengesetzt ist. Dies gilt umso mehr, als berufliche Situationen eine Vielzahl von Fähigkeiten/Fertigkeiten erfordern:

##### - Umgang mit schwierigen Situationen

Misserfolgsmöglichkeiten, Konflikte, Frustrationen, Kritik, usw.

##### - Umgang mit anderen Personen

Persönliche/telefonische Kontakte mit bekannten/ unbekanntem Personen von höherem/gleichen/niedrigerem Rang.

##### - Umgang mit komplexen Leistungsanforderungen

Problemlösung, strukturiertes Vorgehen, Systematik, Planung, Selbstorganisation, Leistungsnormen, etc.

Folgende psychologische Eigenschaften ("Konstrukte") können als persönliche Voraussetzungen für das Zustandekommen von befriedigenden beruflichen Leistungen angesehen werden. Diese Eigenschaften bezeichnen wir normalerweise als „außerfachliche Faktoren der beruflichen Kompetenz“.

- Ausdauer
- Ehrgeiz
- Anspruchsniveau (Gütemaßstäbe für die eigene Leistung, Ansprüche an die eigene sozioökonomische Stellung)
- Schätzung von Erfolgswahrscheinlichkeit bei unterschiedlichen Aufgaben
- Anreizwert von Zielen

- Vermeidung von Misserfolgen
- Selbstwertgefühl, Selbstvertrauen, Selbstbewusstsein
- Extraversion/Intraversion
- "self disclosure" (emotionale Mitteilungen über sich selbst als Basis von Kontakten)
- Emotionale Stabilität
- Leistungsängste
- "Internal/External Locus of Control"  
(die Verursachung von Misserfolg/Erfolg wird in der Person selbst oder in unkontrollierbaren äußeren Umständen gesehen)
- Selbst-Instruktionen, Selbst-Gespräche
- Leistungsmotivation
- und viele andere.

Um solche Konstrukte etwas näher in den Griff zu bekommen, seien zwei davon etwas näher dargestellt.

### Zum Begriff der Leistungsmotivation

Der Begriff der Leistungsmotivation wird in der Literatur je nach dem spezifischen Forschungseinsatz verschieden definiert. Nach GUILFORD (1964) besteht sie aus drei Faktoren, nämlich

- a) allgemeinem Ehrgeiz (general ambition)
- b) beharrlicher Bemühung (persistent effort)
- c) Ausdauer (endurance).

Autoren wie MCCLELLAND (1951) und HECKHAUSEN (1967) definieren die Leistungsmotivation vom Anspruchsniveau, von Normwerten her: Leistungsmotivation als Auseinandersetzung mit einem Gütemaßstab, als Diskrepanz zwischen einem Soll- und einem Ist-Zustand. In Expectancy-Modellen (VROOM, 1964) ergibt sich der Leistungsanreiz aus dem Produkt der Valenz eines Leistungsziels und der Wahrscheinlichkeit der Erreichung dieses Ziels. Die sogenannten Equity-Modelle (COFER & APPLEY, 1964) erklären Leistungsbereitschaft durch einen sozialen Vergleich: die "Bilanz" aus eigenen "Investitionen" (Zeitaufwand, Anstrengung) und eigenem Nutzen (Geld, Status, Anerkennung) im Vergleich zur "Bilanz" von Kollegen. Wenn der Vergleich der eigenen Bilanz mit der der Kollegen in einer Dissonanz endet, dann wird die Leistungsbereitschaft nach oben oder unten geändert. Innerhalb der Verhaltenstheorie schließlich wird die Leistungsmotivation erklärt durch Selbstverstärkung (internal covert self-reinforcement), die auf bestimmte eigene Leistungen gegeben wird. Diese Selbstverwirklichung ist gelernt: Je nach den Einflüssen,

die auf ein Individuum gewirkt haben, wird die Eigenverstärkung nach unterschiedlich hohen Leistungen und in unterschiedlichen Situationen gegeben: bei besonders hart empfundener eigener Anstrengung, bei der Ausschaltung von Konkurrenten, bei der Erledigung der eigenen Arbeit ohne große Anspannung (mit einer gewissen Lässigkeit) und so weiter.

Untersuchungen neuerer motivationstheoretischer Konzepte haben ergeben, dass das Leistungsverhalten abhängig ist vom Grad der erlebten Kontrolle über die Umwelt. Manche Leistungen sind fast voll auf die eigene Begabung, auf den eigenen Einsatz zurückzuführen, manche Leistungen hängen dagegen sehr stark von äußeren Einflüssen ab. Diese "inneren" und „äußeren“ Determinanten der Leistung lassen sich noch in eher stabile und eher variable Einflüsse zerlegen. Auf diese Weise kommt man zu einem Vierfelder-Schema der Leistungsdeterminanten (nach WEINER, 1976)

	äußere Einflüsse	innere Einflüsse
stabile	Aufgabenschwierigkeit Behinderung durch andere (Vorgesetzte, Kollegen)	Begabung Fähigkeiten Interessen
variable	Glück störende oder fördernde momentane Ereignisse	Anstrengung Vorbereitung Indisponiertheit Hemmungen

Ein weiterer Zugang zur Leistungsmotivation besteht zweifellos in den Selbst-Verbalisationen, Selbstbewertungen angesichts von Leistungssituationen. Wie MEICHENBAUM, LURIA, ELLIS und andere Autoren gezeigt haben, wird, durch solche Selbst-Verbalisationen einerseits eine Steuerung des eigenen Handelns hervorgerufen, andererseits erzeugen anscheinend solche Selbst-Verbalisationen Leistungswillen oder Leistungsängste und wirken dabei leistungsdeterminierend. Das vorliegende Messinstrument kann nun so aufgefasst werden, dass damit die entscheidenden Selbst-Verbalisationen erfasst werden und deswegen eine Vorhersage der Leistungsfähigkeit (vom Motivationsgeschehen her) möglich ist.

#### Zum Begriff des Selbstwertgefühls

Das Selbstwertgefühl ist eine psychologische Variable, die in immer stärkerem Maß zur Analyse und Erklärung von Verhaltensweisen herangezogen wird. Dies betrifft sowohl den gesamten sozialpsychologischen Verhaltensbereich wie auch die klinische Psychologie. Im sozialen Bereich wird durchgängig festgestellt, dass Personen mit einem hohen Selbstwertgefühl höheren Erfolg bei anderen Personen haben. Im klinischen Bereich zeigt sich immer wieder, dass psychische Störungen immer eine Störung des Selbstwertgefühls beinhalten. In einigen Fällen dürfte es sogar so sein, dass eine massive Störung des Selbstwertgefühls andere psychische Störungen nach sich ziehen werden.

Bei der Definition des „Selbstwertgefühls“ soll hier einem Ansatz von SECORD (1963) gefolgt werden, der vorschlägt, den etwas schwierigen Begriff des "Selbst" zu

zerlegen in die sogenannten "self-referent-behaviors". Diese Verhaltensweisen sind also dadurch charakterisiert, dass eine Person sich selbst als Objekt gegenübersteht und sich diesem Objekt gegenüber in einer bestimmten Weise verhält. Das Selbstwertgefühl wären demnach also Bewertungen der eigenen Person, der eigenen Leistungen, des eigenen Aussehens, der eigenen Kompetenz, und so weiter. Das Selbstwertgefühl ist damit auch eine Art von cognitivem Schema, das die Basis für Einschätzungen des Erfolges oder Misserfolges in bestimmten Situationen abgibt. (PIONTKOWSKI 1975).

Im sozialpsychologischen Bereich haben sich folgende Korrelate von hohem bzw. niedrigem Selbstwertgefühl gefunden:

- Personen mit hohem SWG tendieren in stärkerem Maß dazu, Einfluß auf andere auszuüben, als Personen mit niedrigerem SWG (COHEN, 1959)
- Personen mit hohem SWG sind kreativer und fühlen sich auch in bislang unbekanntem Situationen kompetent. Personen mit hohem SWG zeigen weniger Widerstand gegen organisatorische Veränderungen.
- Personen mit hohem SWG lassen sich nicht so leicht beeinflussen wie Personen mit niedrigem SWG. Sie zeigen mehr Vertrauen in die eigene Meinung.
- Personen mit hohem SWG haben eine positivere Einstellung gegenüber ihren Interaktionspartnern.
- Personen mit niedrigem SWG haben größere Schwierigkeiten, ihre Bedürfnisse und Aggressionen auszudrücken. Sie sind vor allem besorgt darum, was andere Leute von ihnen denken.
- Personen mit niedrigem SWG, sind in ihren Gefühlen leicht verletzbar, sind gegenüber den Personen, die sie mögen, öfter ärgerlich oder gereizt.

- Personen mit hohem SWG sind sozial erfolgreicher und beliebter als Personen mit niedrigem SWG (ZILLER et al., 1969).

Im klinischen Bereich wird ein positives SWG geradezu als Indikator für psychische Gesundheit, ein niedriges SWG als Indikator für psychische Störungen angesehen. Wie MEICHENBAUM (1972) mit großer Deutlichkeit darstellte, neigen Personen mit einem niedrigen SWG zu bestimmten Arten von "Selbstinstruktionen", zu bestimmten Eigensteuerungen des Verhaltens, das sozusagen den Misserfolg schon vorprogrammiert. SATIR (1975) beschreibt, wie ein niedriges SWG zu bestimmten interpersonellen Verhaltensweisen führt, die in keiner Weise geeignet für die Konfliktbehebung sind. Diese Verhaltensweisen sind beschwichtigen, anklagen, rationalisieren und ablenken. Durch diese Verhaltensweisen werden Konflikte nicht gelöst, sondern entweder zementiert oder verschärft.

Durch diese kurze Darstellung von Korrelaten des hohen und niedrigen SWG zeigt sich, dass diese Variable äußerst wertvoll für die Vorhersage des Verhaltens werden kann. Ausführliche Darstellung der Zusammenhänge: Siehe STRASSER, 1975.

## 7.5. Validität der Messwerte

Eine Eigenschaft von psychologischen Meßmethoden ist, dass man eigentlich nicht genau weiß, was sie erfassen. Selbst ein so einfaches Instrument wie ein Fragebogen zur Leistungsmotivation ist in Wirklichkeit eine hoch-komplizierte Angelegenheit. Es geht darum, herauszufinden, wie viel Prozent der Varianz eines Testwertes wirklich auf Unterschiede in der betreffenden Eigenschaft zurückzuführen ist. In einem Schaubild von möglichen Varianzanteilen ausgedrückt:



Betrachtet man sich dieses Diagramm, dann wird klar, dass Validitätskoeffizienten zwangsläufig in der Regel relativ niedrig ausfallen müssen, da meist mehrere irrelevante Varianzanteile in ein Testergebnis mit einfließen.

Über den genauen Zusammenhang zwischen diesen Einflüssen und dem Testwert habe ich an anderer Stelle berichtet.

(STRASSER, Über die Validität von Persönlichkeitsfragebögen und biografischen Informationsbögen, Dissertation, 1975, Univ. München).

Deshalb sollten hier nur einige Punkte herausgegriffen werden.

### 7.5.1. Situative Einflüsse

Tests/Fragebögen insbesondere, werden in der Regel anders beantwortet, ob man sich in einer Bewerbungssituation befindet oder nicht. In der Bewerbungssituation spielt der Einfluss der "social desirability" natürlich eine wesentlich stärkere Rolle. Die genaue Größe des Effekts und das Ausmaß von social desirability-Verschiebungen der Testwerte bei unterschiedlichen Persönlichkeitstypen ist derzeit nicht genau bekannt.

### 7.5.2. Zufällige Fehlantworten

Ein wichtiges Kriterium für ein Testsystem ist die Frage, ob und in welchem Umfang zufällige Fehlantworten als Fehlervarianz des Testwertes ausgeschlossen werden kann.

Dieses Kriterium kann überprüft werden, indem man die Häufigkeit von sozial unerwünschten Antworten (bei extrem formulierten Items) ermittelt. Zu diesem Zweck wurde bei zwei Stichproben (Frankreich N=47 und England N=57) die Beantwortung von extrem formulierten Items überprüft.

Zufallsfehler:

Beantwortung extrem "unerwünschter" Alternativen

	Item-Nr.	England	Frankreich
1	11	0	0
2	16	0	2
3	18	0	0
4	32	0	0
5	50	0	2
6	67	0	0
7	68	0	0
8	72	0	2
9	78	0	2
10	84	0	0
11	95	0	0
12	101	0	0
13	116	0	0
14	130	0	0
15	131	0	2
16	135	0	0
17	136	2	0
18	140	0	2
19	141	0	2
20	145	0	0

Wie aus der Tabelle hervorgeht, scheint eine zufällige Fehlantwort äußerst unwahrscheinlich zu sein; so dass eine entsprechende Fehlervarianz vernachlässigt werden kann.

Bei der Interpretation der Tabelle ist zu berücksichtigen, dass der Ausschluss der Alternativen ja nicht 100 % ist, sondern einige Testpersonen durchaus die entsprechende Alternative ankreuzen könnten.

### 7.5.3. Item-Schwierigkeiten und Trennschärfen

Einzelne Items sollten zwischen Personen mit hohen und niedrigen Werten auf einer Skala identifizieren. Wenn in einer Zielpopulation z.B. bestimmte Items nicht über die unterschiedlichen Antwort-Alternativen "streuen", dann kann natürlich kein Beitrag zur Differenzierung auf der gesamten Skala erwartet werden. Dies wird normalerweise mit den Indices "Itemschwierigkeit" und "Trennschärfe" errechnet.

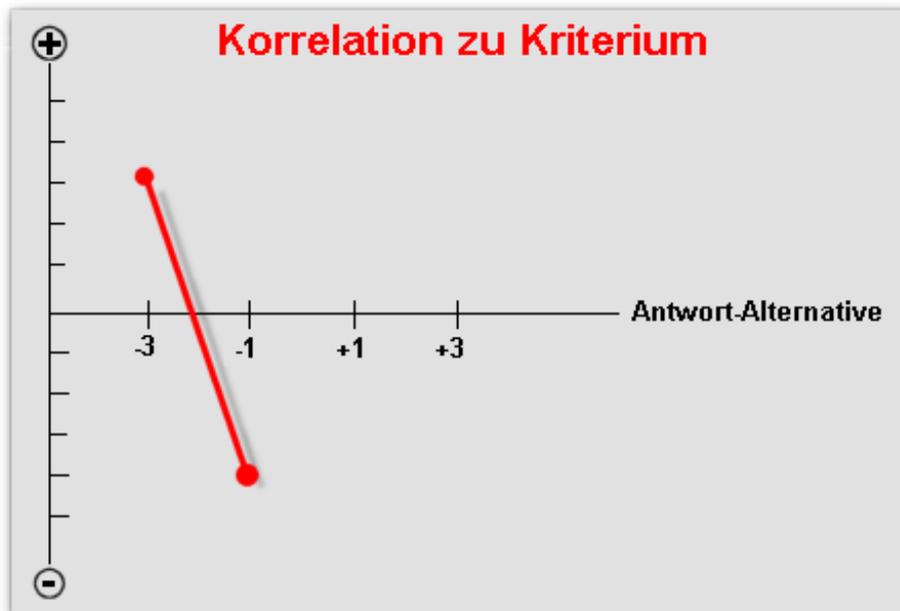
Anstelle von Itemschwierigkeit wurden Verteilungen auf die einzelnen Antwortalternativen bei jedem Item ermittelt. Dies schien uns ein genaueres Bild zu geben als die Größe "Itemschwierigkeit". (Siehe Anhang)

Bei der Überprüfung der Antwortenverteilung auf die einzelnen Alternativen fällt auf (siehe Tabelle), dass es sowohl normal verteilte wie stark dichotom verteilte Alternativen gibt.

Dichotom verteilte Items sind vor allem diejenigen, die stark positiv oder negativ formuliert sind.

Bei einem negativ formulierten Item kommt eine Bejahung so gut wie nicht vor; umgekehrt gilt das auch für positiv formulierte Items und einer Verneinung. Aber offenbar sind gerade die dichotom verteilten Items sehr valide, weil die völlige Verneinung (--) hoch mit Erfolg, die leichte Verneinung (-) hoch mit Misserfolg korreliert. Im umgekehrten Sinn gilt dies auch für die positiv getexteten Items.

Schematische Darstellung des Korrelations-Verlaufs bei dichotomen Items



## Trennschärfen von Items

Beispiel:      Faktor AMS

94	0,51
102	0,44
87	0,58
39	0,31
75	0,54

Trennschärfe-Indices für weitere Faktoren finden sich bei SEDLMAYR (1969).  
Insgesamt sollten die Trennschärfen mehrmals überprüft werden; die Daten für die Hauptfaktoren des Tests zeigen jedoch sehr zufriedenstellende Werte.

#### 7.5.4 Trainierbarkeit von Tests als intervenierende Variable

Für die Aussagefähigkeit eines Tests wird in letzter Zeit immer mehr die "Trainierbarkeit zum Kriterium.

Die Vorbereitung auf Bewerbungssituationen ist mittlerweile ein Thema von zahlreichen Büchern sowie Seminaren. Dort werden "Knack-Möglichkeiten" für die gängigsten Tests (z.B. IST, 16 PF) trainiert.

Damit schwindet jede Aussagefähigkeit, weil der Hauptunterschied sich zwischen "trainierten" und "naiven" Bewerbern ergeben dürfte.

Für eine wirklich valide Aussage ist also Nicht-Bekanntheit eines Tests bzw. Neuheit der Item-Gestaltung ein wichtiges Kriterium.

#### 7.5.5. Stabilität/Instabilität von Eigenschaften

Eine wichtige Voraussetzung der Messung von psychischen Eigenschaften ist die Frage nach der grundsätzlichen Stabilität dieser Eigenschaften. Manche Theorien behaupten, dass es eigentlich keine stabilen Eigenschaften gäbe, dass Selbstvertrauen, Leistungsmotivation, Stabilität, etc. völlig mit den jeweiligen Situationen variierten.

Dies ist im extremen Ausmaß sicherlich nicht so, denn dann würden sich Menschen völlig vorhersagbar verhalten, was offensichtlich nicht der Fall ist. Eine gewisse situative Variabilität ist aber vorhanden. Einen Teil dieser Instabilität nennen wir "Oszillation". Der Grad der Oszillation ist nicht ein Beitrag zur Fehlervarianz eines Testwerts, sondern ein wichtiges diagnostisches Kriterium.

Sehen wir uns das Beispiel "**Extraversion**" (Kontaktfreude) an. Für viele Menschen gilt, daß sie teilweise kontaktfreudig, teils aber auch manchmal gerne allein sind. Das Verhalten ist also schwankend, "oszillierend". Erfasst man nur einen einzigen Wert, dann weiß man eigentlich nicht genau, worum es sich handelt. Es ist notwendig, die Stärke der jeweiligen Extraversions- und Rückzugsanteile zu erfassen und auch die Stärke der Oszillation. Wenn eine Person stark oszilliert, wird sie unberechenbar und frustrierend im Kontakt, man wendet sich eher von ihr ab.

**Für die Test-Statistiker:** Wenn man nur einen Wert des Extraversionsbereichs mit dem Führungserfolg validiert, bekommt man zwangsläufig niedrigere Validitäten, als man bei Berücksichtigung der Oszillation bekommen würde.

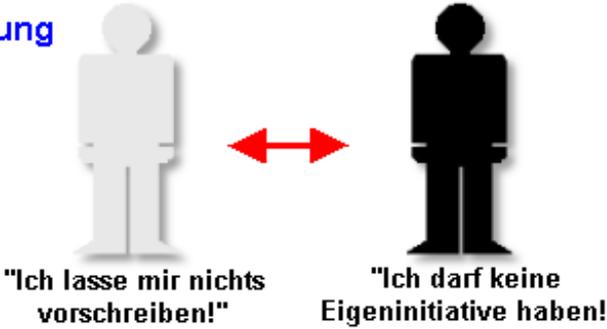
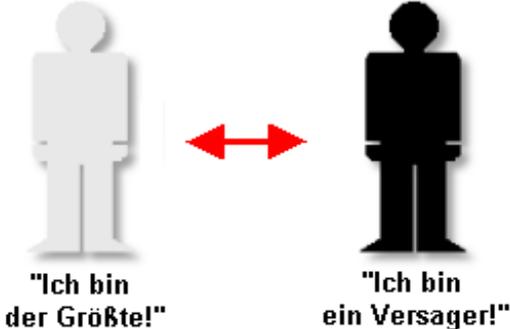
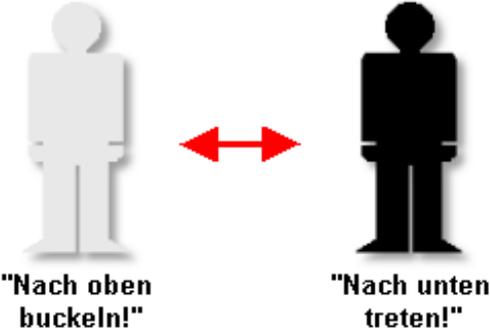
## Oszillation

Wenn man Persönlichkeitseigenschaften in Beziehung zu Bedürfnisstrukturen setzt, dann ergibt sich folgende theoretische Erklärung von Oszillation:

Durch Lebenserfahrungen werden Bedürfnis- und Motivationsstrukturen modifiziert, "verschoben".

Die Verschiebung von Bedürfnissen geht meist in eine bestimmte Richtung (entweder zu sehr positiv oder zu negativ). Das bedeutet, dass ein Ungleichgewicht entsteht und damit ein Gegenpol. Personen mit einem übersteigerten Selbstwert-Bedürfnis haben deshalb gleichzeitig Größenideen auf der einen Seite und latente Minderwertigkeitskomplexe auf der anderen Seite. Dieser Vorgang heißt "Oszillation"; er bedeutet, dass Ausgewogenheit, dass die Mitte fehlt und die Person von einem Extrem ins andere pendelt. Meist ist es so, dass ein Pol im Vordergrund des Verhaltens steht und der andere Pol nur gelegentlich durchbricht. Dieser verdrängte Pol wird auch "Schatten" genannt. Wenn eine ausgeprägte Oszillation vorliegt, dann wird das Verhalten des Betroffenen sehr schlecht einschätzbar und vorhersagbar; man weiß nie genau, woran man ist. Auch der Betroffene selbst fühlt sich zwiespältig, zerrissen, gespalten; an einem Tag möchte er die Welt umarmen und "in der Menge baden", am nächsten Tag zieht er sich total zurück und will mit niemandem etwas zu tun haben.

## EINIGE TYPISCHE OSZILLATIONEN

Bedürfnis	Oszillation	
<b>Eigenbestimmung</b>	 <p>"Ich lasse mir nichts vorschreiben!"</p> <p>"Ich darf keine Eigeninitiative haben!"</p>	Schwierige Menschen, die zwischen Mangel an Eigeninitiative und Kampf gegen alle Fremdbeeinflussung schwanken und oft große Arbeits- und Leistungsstörungen aufweisen.
<b>Selbstwertgefühl</b>	 <p>"Ich bin der Größte!"</p> <p>"Ich bin ein Versager!"</p>	Schwanken zwischen Größenwahn und Minderwertigkeitskomplexen. Große Pläne werden geschmiedet, die Phantasie läuft auf Hochtouren, nur in der Praxis geschieht selten etwas Nennenswertes.
<b>Dominanz</b>	 <p>"Nach oben buckeln!"</p> <p>"Nach unten treten!"</p>	"Radfahrer-Verhalten"
<b>Gerechtigkeit</b>	 <p>"Ich opfere mich für Dich!"</p> <p>"Aber Du musst Dich auch total für mich opfern!"</p>	Einerseits Aufopferung, andererseits (unausgesprochene) Riesenerwartungen an den anderen.

**Bedürfnis**

**Oszillation**

**Geselligkeit**



**"Ich will Geborgenheit!"**



**"Man mag mich sowieso nicht!"**

Große Schwankungen, abrupter Wechsel zwischen Suche nach Zuwendung und Ablehnung anderer Menschen. Hohe Eifersucht, große Angst vor dem Allein-Sein.

**Sicherheit**



**"Mir kann nichts passieren."**



**"Ich habe Angst."**

Wechsel zwischen tollkühnem Verhalten und Feigheit, Verzagtheit.

Diese Oszillationen werden auch beschrieben von

HORNEY (1945) , FROMM (1937) , FROMM (1980) sowie RICHTER (1979).

### 7.5.6. Situative Instabilität der Testwerte

Ein wichtiger situativer Einfluss auf die Testergebnisse scheint der Grad der persönlichen Stressbelastung und Zufriedenheit mit den Lebens-/Arbeitsumständen zu sein.

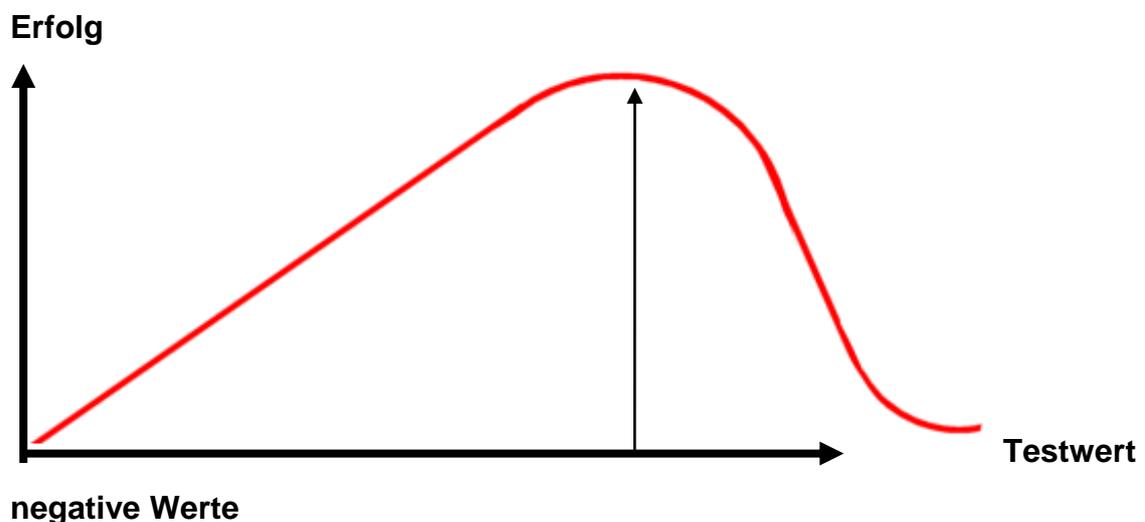
Es wurde immer wieder gefunden, dass bei hoher Stressbelastung bzw. großer Unzufriedenheit mit der persönlichen Situation die Testwerte deutlich negativer ausfallen als bei allgemeiner Harmonie.

In seiner derzeitigen Version erlaubt das Verfahren noch keine systematische Korrektur dieser Einflüsse. Diese Effekte können im zugehörigen Gespräch jedoch herausgefunden und in das Gesamturteil mit einbezogen werden.

### 7.5.7. Merkmalsausprägung im Berufserfolg

Die Voraussage des Berufserfolgs beruht im wesentlichen auf der Korrelationsrechnung, also auf linearen Methoden: je mehr einer Eigenschaft, um so höher der wahrscheinliche Berufserfolg. Diese linearen Modelle stimmen mit der Wirklichkeit offenbar nicht gut überein, was zu finden ist, sind im wesentlichen kurvilineare Zusammenhänge zum Erfolg. Diese Zusammenhänge haben im allgemeinen die folgende Form:

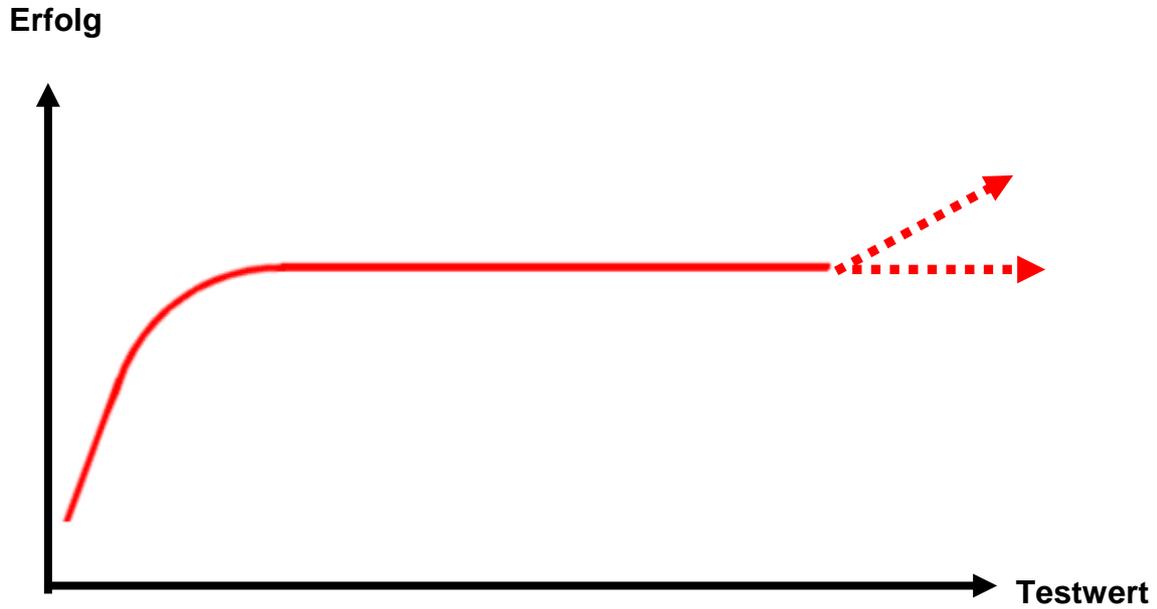
## Übersteigerungen



Das bedeutet zum Beispiel: Hohe Kontaktfreude ist gut, aber zu hohe Kontaktfreude kann negativ für den Erfolg sein. Systematik ist gut, aber zu viel Systematik ist hinderlich. Das gilt für (fast) alle Faktoren.

***Für die Teststatistiker:*** Diese Kurvilinearität beeinträchtigt natürlich stark die Validität eines Tests, wenn man mit linearen Regressionen arbeitet. Leider sind Methoden der kurvilinearen Regression extrem kompliziert; so dass in keinem Bereich (am allerwenigsten bei Interviews oder Assessments) damit gearbeitet wird.

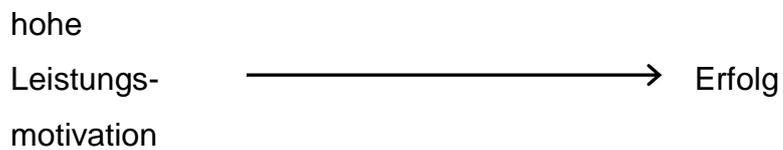
## Indifferenz - Bereich



Das bedeutet: abgesehen von Extrem-Ausprägungen besteht keine Beziehung zu Kriterien. Beziehungen dieses Typs werden oft bei Analysen der Leistungsmotivation gefunden: nur in Extremgruppen negativer, oder positiver Motivation ergeben sich Beziehungen zur tatsächlichen Leistung. (SEDLMAYR, 1969)

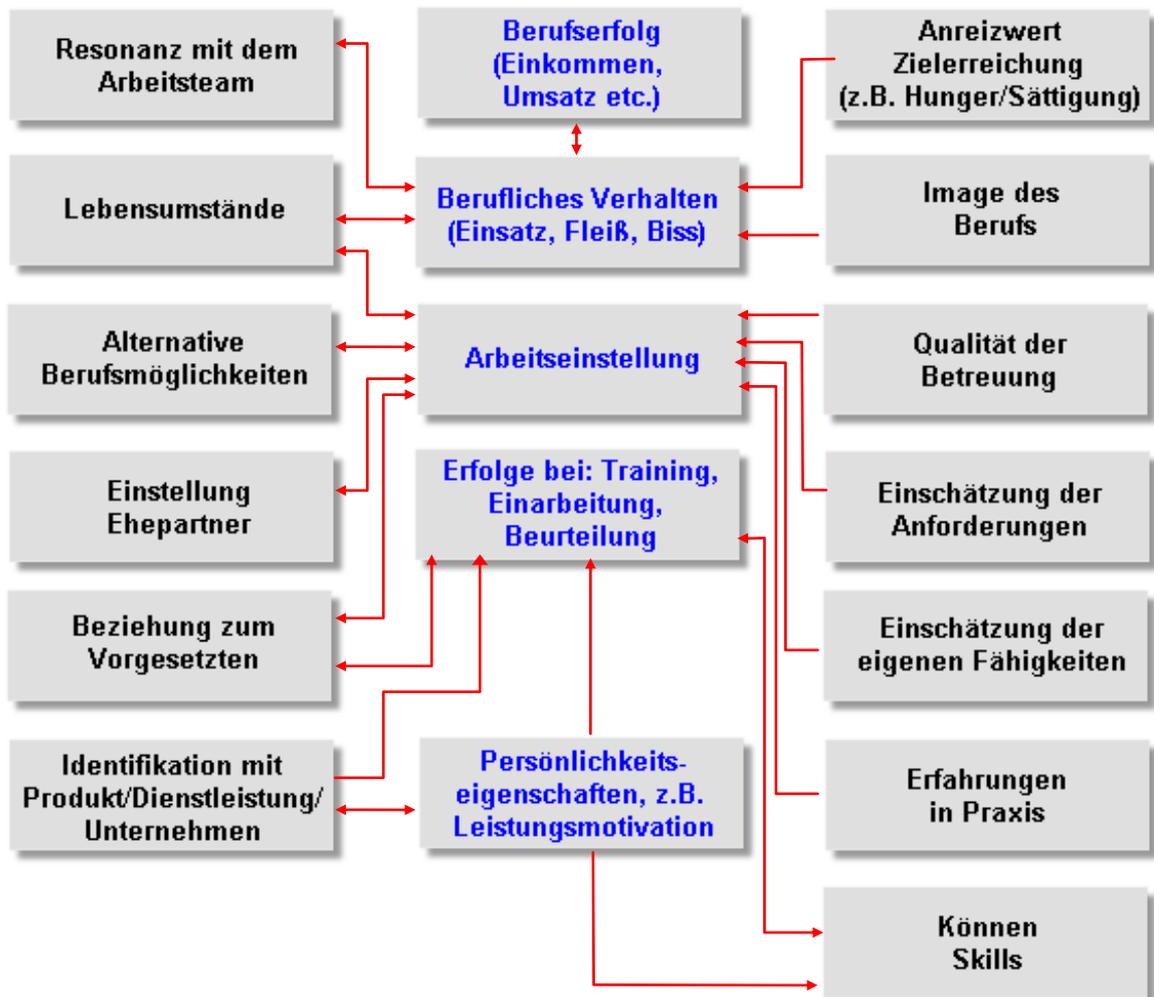
## 7.6. Determinanten des Berufserfolgs

Wenn man den Berufserfolg vorhersagen will, dann sollte man plausible Annahmen darüber haben, was diesen Erfolg ausmacht. Meist geht man von einfachen Vorstellungen des Zusammenhangs aus, etwa der folgenden Art:



Diese einfachen Beziehungen sind in der Regel aber nicht trag-fähig. Nach unseren bisherigen Erfahrungen sollte man von einem komplexeren Modell ausgehen, in dem die folgenden Komponenten Eingang finden:

Zusammenhang zwischen Berufserfolg  
und psychischen Eigenschaften  
(Varianzanteile der Erfolgsbedingungen)



Dieses Diagramm weist mit aller Deutlichkeit auf einen Umstand hin: Der Berufserfolg kann niemals mit einem Test allein vorhergesagt werden. Alle anderen Einflüsse würden dabei unter den Tisch fallen, was die Treffsicherheit sicherlich untergraben würde. Andererseits ist es offenbar so, dass der Verzicht auf Tests auch nicht realistisch ist, da eine Vorhersage des Berufserfolgs nur anhand von schriftlichen Unterlagen oder Interviews sehr geringe Validitäten aufweist. (Dies wird später noch näher dargestellt.)

Die ideale Kombination scheint Test + Interview zu sein.

Zweckmäßigerweise sollte man deshalb nicht von der Validität eines isolierten Tests ausgehen, sondern von der Validität der kombinierten Auswahlprozedur Test + Interview.

Dies gilt umso mehr, als unserer Erfahrung nach wesentlich aussagekräftigere Interviews durchgeführt werden, wenn die Testergebnisse mit einbezogen werden.

Aus Erfahrungsberichten wissen wir, dass es Misserfolge selbst dann geben kann, wenn anscheinend alle Voraussetzungen stimmen: qualifizierte Personen werden qualifiziert ausgebildet und kommen zu qualifizierten Verkaufsleitern, und trotzdem kommt kein Erfolg zustande. Dies liegt an vielen Einflussmöglichkeiten, die prinzipiell in diesem Schema enthalten sind. Von besonderer Wichtigkeit für den Erfolg scheint die Beziehung zum Vorgesetzten zu sein. Man kann fast sagen, dass die eigentliche Qualifikation (Extremwerte ausgenommen) lange nicht so wichtig ist wie ein positives Resonanzfeld mit dem Vorgesetzten. Die Dynamik zwischen dem neuen Mitarbeiter und seinem Vorgesetzten ist in folgendem Schema niedergelegt:

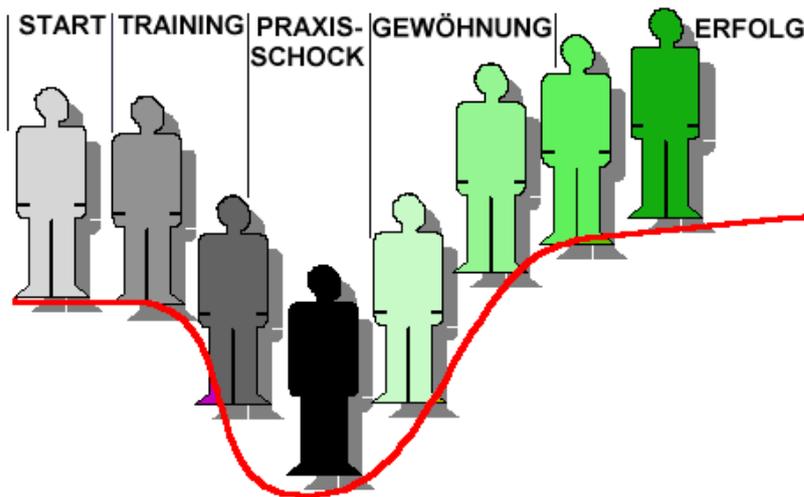
**Dynamik des Zusammenhangs zwischen Erfolg und Beziehung zum Vorgesetzten**

		<b>Hohe Sympathie</b> Stärken und Schwächen im Bereich sozialer Kompetenz sind bekannt.	<b>Niedrige Sympathie</b> Stärken und Schwächen im Bereich sozialer Kompetenz sind nicht bekannt.
<b>Verhalten des Mitarbeiters</b>	<b>positiv</b>	<b>"Ich wußte es, das ist der Richtige!"</b> (Positive Reaktion, Verstärkung des positiven Verhaltens)	<b>"Das war sicher nur Zufall."</b> (keine Anerkennung, Verhalten des Mitarbeiters wird nicht gefördert)
	<b>negativ</b>	<b>"Ist nicht so schlimm, wird schon werden. Ich weiß wo ich helfen (fördern) muß!"</b> (keine Reaktion, Löschung des Fehlverhaltens)	<b>"Ich wußte es gleich, das wird nichts."</b> (Kritik; Abbau des Selbstvertrauens beim neuen Mitarbeiter, Erhöhung von Fehlreaktionen)

Qualität der Einarbeitung eines neuen Mitarbeiters als intervenierende Variable bei der Validität und der Beziehung zum Vorgesetzten

Offensichtlich spielt für den Berufserfolg eines neuen Mitarbeiters neben seinen Qualifikationen auch die Qualität der Einarbeitung eine entscheidende Rolle.

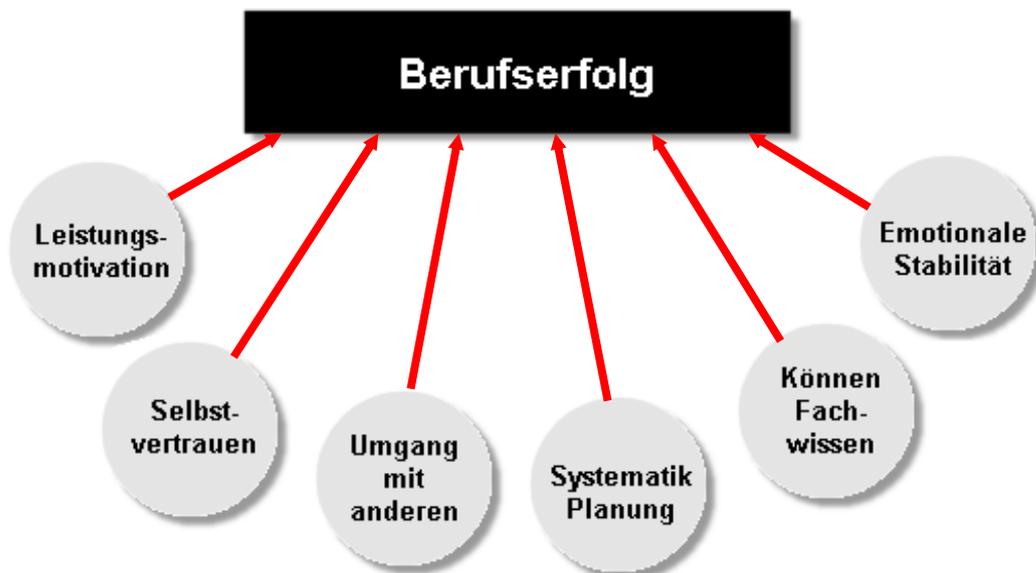
Eine sehr wichtige Variable bei der Einarbeitung neuer Mitarbeiter, insbesondere in schwierigen Verkaufsberufen, ist der Motivationsverlauf, der normalerweise in den folgenden Phasen abläuft:



## 7.7. Vollständigkeit der Erfassung wichtiger Eigenschaften

Die kurze Beschreibung der Bedingungsfaktoren des Berufserfolges zeigt, dass vermutlich viele Einzelfaktoren den Erfolg determinieren. Bei entsprechender Analyse haben wir gefunden, dass es nicht z.B. das Selbstvertrauen allein sein kann, das über Erfolg und Misserfolg unterscheidet. Es gibt Erfolgreiche mit hohem aber auch mit niedrigem Selbstwertgefühl. Diejenigen Personen mit niedrigem Selbstwertgefühl kompensieren dieses aber mit Stärken in anderen Bereichen wie z.B. erhöhter Einsatz, Fachwissen, Zuverlässigkeit. Nur die Berücksichtigung vieler Faktoren, die in Zusammenhang mit dem Berufserfolg stehen, kann deshalb einen vernünftigen Beitrag zur Auswahlentscheidung liefern.

Berufserfolg ist eine Funktion vieler Vorbedingungen:



Eine gute Voraussage ist nur durch Berücksichtigung vieler Faktoren möglich.

Diese Erkenntnis war ausschlaggebend für die Zusammenstellung des vorliegenden Tests. In ihm wurden Faktoren, die sich in unterschiedlichen Situationen bewährt hatten, zusammengestellt, um eine möglichst breite Messung von Vorbedingungen des Berufserfolges zu gestalten. Die Bewährung der einzelnen Faktoren als valides Messinstrument ist in den Ergebnissen der Validierungs-Situation niedergelegt.

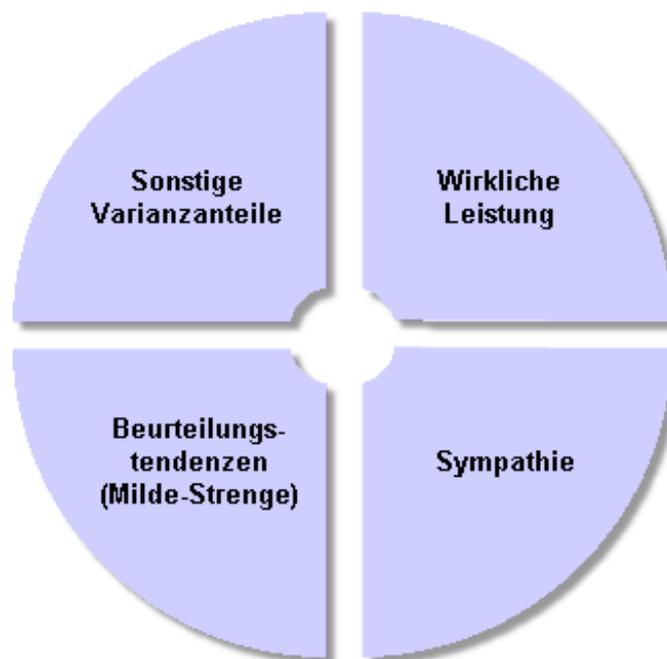
Die Tatsache, dass Berufserfolg ein Ergebnis des Vorhandenseins vieler Faktoren ist, führt auch noch zu einer Korrektur der Einschätzung der Höhe von Validitäts-Koeffizienten. Angenommen, der Berufserfolg ist eine Resultierende von insgesamt 10 Eigenschaften, wie z.B. Fachkenntnisse, Motivation, Stabilität, Extraversion, etc. Dann würden auf jeden Faktor ca. 10 % gemeinsame Varianz mit dem Kriterium entfallen, was einem Validitätskoeffizienten von maximal 0,35 entspricht.

Ein sehr häufiger Fehler in der Anwendung von Tests ist, dass man sich auf ein Verfahren verlässt und dessen Ergebnisse überinterpretiert. Ein Test kann durchaus valide sein, der Beitrag zur Erklärung der Varianz des Erfolgskriteriums beträgt aber nur einige Prozent. Eine Überinterpretation erfolgt deswegen, weil mangels anderer quantitativer Kriterien oft nach diesem "Strohalm" zur Begründung von Entscheidungen gegriffen wird. Aus diesem Dilemma entkommt man nur, indem man versucht, möglichst vollständig die Kompetenzfaktoren zu erfassen, die für den Erfolg in einer bestimmten Position notwendig sind.

## 7.8. Was ist überhaupt Erfolg?

Wenn man den Erfolg voraussagen will, dann sollte man sich darüber im klaren sein, welche Kriterien des Erfolgs vorhanden sind. Hier gehen Schwierigkeiten los, denn Erfolgskriterien sind teilweise nicht vorhanden oder nur zum Teil brauchbar.

Nehmen wir als Erfolgskriterium die Einschätzung durch den Vorgesetzten. Diese Einschätzungen sind wahrscheinlich aus den folgenden Varianz-Anteilen zusammengesetzt:



Auch der Umsatz kann nicht als alleiniges Erfolgskriterium herangezogen werden, da er ebenfalls von vielen Einflüssen abhängt, die nichts mit der eigentlichen Leistung des Außendienst-Mitarbeiters zu tun haben.

Auch im Führungsbereich stößt man bei der Beurteilung des Erfolgs auf Schwierigkeiten. Die Beförderung zum Vorgesetzten hängt nicht immer mit der Leistung zusammen:

Oft werden Kompromisse eingegangen.

Zwei Gruppierungen innerhalb einer Organisation einigen sich auf einen schwachen Kandidaten, der keinem Lager gefährlich werden kann.

Beziehungen geben den Ausschlag.

Oder man einigt sich auf eine Person, die zwar hohe Fachkenntnisse und langjährige Erfahrung, aber deutliche Lücken in anderen Kompetenzbereichen aufweist.

Das Glück, das gerade jemand dringend benötigt wurde, hat schon vielen zu Positionen verholfen, denen sie nicht gewachsen waren.

Oft erfolgt Beförderung auch nur, um langjährigen verdienten Mitarbeitern eine "Belohnung" zu geben.

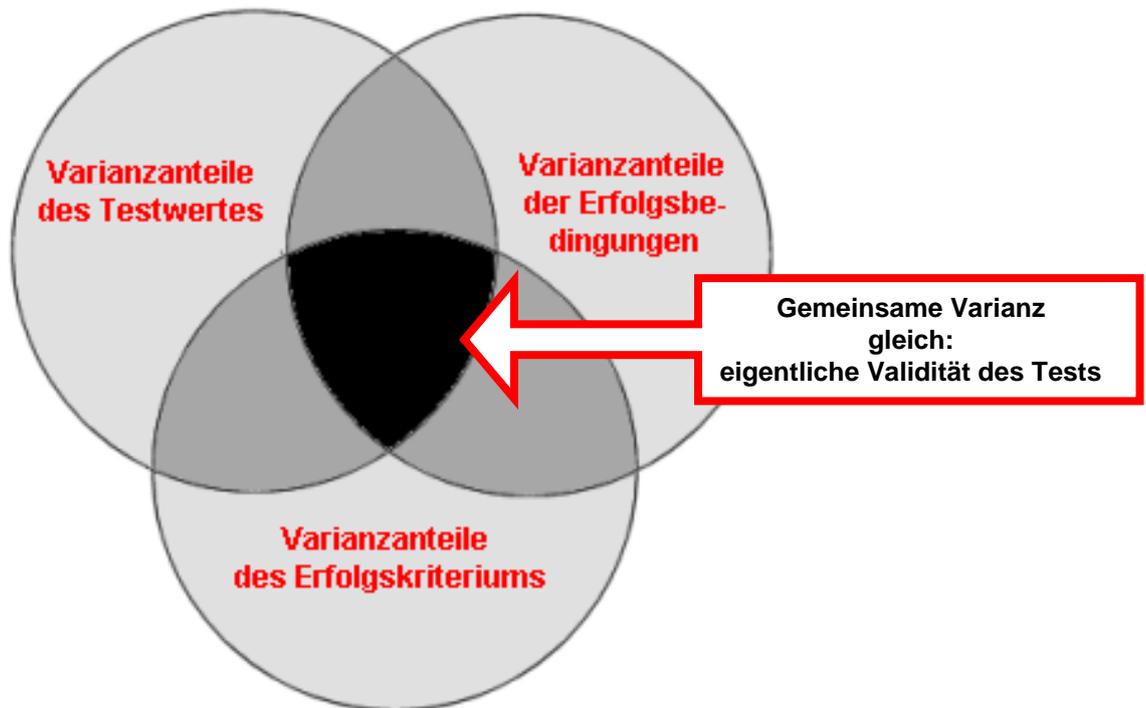
Diese irrelevanten Varianzanteile mindern die Höhe der Validitätskoeffizienten.

### 7.9. Welche Validitäten sind realistisch?

Oft wird beklagt, dass die Validitäten von Tests zu niedrig sind. Auch wenn sie im Bereich von statistischer Signifikanz liegen (das bedeutet, dass die Wahrscheinlichkeit, dass der Koeffizient nur auf Zufall beruht, gegen 0 geht), dann bewegen sie sich in Bereichen von 0,20 bis 0,40 - das sei schließlich zu wenig.

Falsch. Das sind genau die Werte, die zu erwarten sind. Höhere Werte für einzelne Testfaktoren können nicht erreicht werden. Wenn sie trotzdem vorhanden sind, beruhen sie auf Zufällen.

Wenn wir das bislang Besprochene nochmals in einer Grafik zusammenfassen, dann ergibt sich folgendes Bild:



Das bedeutet: die rein numerischen Validitäten müssen für einen Test zwangsläufig relativ niedrig sein. Alles was (für einen einzelnen Faktor) über 0,35 hinausgeht, dürfte eher zufälligen Charakter haben.

Schlechte Leistungen können sicherer vorhergesagt werden als gute Leistungen.

Dies hängt offenbar damit zusammen, dass es viele unterschiedliche Wege zum Erfolg gibt, die nur zum größten Teil, aber nicht vollständig durch den Test erfasst werden.

Zum Beispiel gibt es viele unterschiedliche, zum Teil sehr individuelle Wege zum Erfolg als Verkäufer. Der Test kann aber nur die häufig verwirklichten "Standard-Voraussetzungen" erfassen, nicht aber die Ausnahmen.

## Nicht-lineare Validitätsbeziehungen

In einem Vier-Felder-Schema ausgedrückt:

		Leistung	
		+	-
Testwert	+	o.k.	Fehler 1
	-	Fehler 2	o.k.

Fehler 1: "Blender-Effekt":  
Gute Testwerte, aber schlechte Leistungen.  
Gute Fähigkeiten, aber (derzeit) nicht motiviert.

Fehler 2: Sehr komplexer Natur:  
Möglichkeiten:  
- gutes Potential, aber situative Frustration  
- Überkompensation von persönlichen Schwierigkeiten durch die Arbeit  
- untypische Erfolgswege  
- Erfolg aufgrund langjähriger guter Beziehungen, nicht aufgrund derzeitiger höherer Leistungsfähigkeit

In der Praxis heißt das: Durch das Testergebnis werden gelegentlich "Hochleister" zu unrecht abgelehnt.

Für diese Validitätskoeffizienten bedeuten diese Zusammenhänge natürlich eine weitere Einbuße der numerischen Höhe.

## 8. Vergleich: Validitäten von anderen Mess-Methoden

Um zu einer realistischen Einschätzung der gefundenen Validitätskoeffizienten des Tests zu kommen, ist es sicherlich hilfreich, sich Validitäten von anderen Methoden zu Voraussagen des Berufserfolgs anzusehen.

## **BRAUCHBARKEIT VON BEWERBUNGSUNTERLAGEN/ZEUGNISSEN**

- \* soweit frühere und künftige Berufstätigkeit sich nicht in hohem Maß gleichen, ist die prognostische Validität der Schlüsse aus Bewerbungsunterlagen recht gering:

$$r = 0.14$$

(10 Untersuchungen)

- \* Validität von Noten (Schulnoten/Examenszeugnisse)

- für Ausbildungs-Noten: 0.40 - 0.50

- für den allgemeinen Berufserfolg: 0.1 5

*Quelle:* SCHULER, H. Berufseignungsdiagnostik.

In: Enzyklopädie der Psychologie,

Organisations-Psychologie, Göttingen (HOGREFE)

1989

## **BRAUCHBARKEIT VON INTERVIEWS**

Das Interview wird von 56 % von befragten Unternehmen als das wichtigste Selektionskriterium bezeichnet. Nur 2% hatten allerdings die Validität des Interviews überprüft. Bei keiner anderen Methode ist die Diskrepanz zwischen subjektiver Wertschätzung und empirischer Bewährung so ausgeprägt wie beim Interview.

Reliabilität des Interviews: 0.50

Validität: zwischen 0.08 und 0.20, wahrscheinlichster Wert: 0.14  
(Ergebnisse von mehr als 20 Einzelstudien)

*Quelle:* SCHULER, H. Berufseignungsdiagnostik.

In: Enzyklopädie der Psychologie, Organisationspsychologie,  
Göttingen (HOGREFE), 1989

## **BRAUCHBARKEIT VON ASSESSMENT CENTERS**

Validität: 0.37 - 0.45

(Ergebnisse aus mehr als 70 Einzelstudien)

Validität im besonderen dann hoch, wenn im Assessment  
psychologische Tests mitverwendet werden.

*Quelle:* SCHULER, H. Berufseignungsdiagnostik. In:  
Enzyklopädie der Psychologie, Organisationspsychologie,  
Göttingen (HOGREFE) 1989

## 9. Validierungsstudien

## 9.1. Ergebnisse von Validierungsstudien

Da Erfolgskriterien, wie bereits dargestellt, erstens von einer gewissen eigenen Fehlervarianz behaftet sind und zweitens in der Regel sehr stark "kondensierte" Indices sind, wurden im Lauf der Zeit Untersuchungen mit mehreren sehr unterschiedlichen Außenkriterien angefertigt. Diese Kriterien waren dabei im Wesentlichen:

1. Zahl und Dauer von Abwesenheit/Krankheitstagen
2. Ergebnisse von Leistungstests
3. Verkaufserfolge
4. Berufserfolge
5. Schulnoten
6. Verhaltensanalysen
7. Externe Gruppenvergleiche
8. Beurteilung durch Vorgesetzte

### Zusammenfassung:

In allen Fällen wurden signifikante Zusammenhänge zwischen den Faktoren des DNLA-Verfahrens (bzw. Vorläufern oder vergleichbaren Faktoren) und den Kriterien gefunden. Die Höhe der Korrelationen bewegen sich in der Regel zwischen 0,25 bis 0,60.

## 9.2. Validität zu Krankheits-/Abwesenheitsdauer/-tagen

(N = 186)

Die Faktoren NAA und UNN zeigten signifikante Korrelationen zu einem kombinierten Index aus Abwesenheitshäufigkeit und -dauer in verschiedenen untersuchten Firmen. Es ist klar, dass diese beiden Indices von einer ganzen Reihe von Rahmenbedingungen beeinflusst werden, die nichts mit Persönlichkeitsfaktoren zu tun haben, wie z.B. die allgemeine konjunkturelle Lage usw. Aus diesem Grund sind hier signifikante Korrelationen als sehr aussagekräftig zu beurteilen.

## 9.3. Beziehungen zu Leistungstests

In einer anderen Untersuchung zeigte sich, dass die verwendeten Faktoren sogar bei reinen Leistungstests (dem Figur-Rekonstruktionstest, BRENGELMANN) valide waren (siehe STRASSER, 1975) . Die erreichten Koeffizienten lagen zwischen .27 und .54, mit einem Median von .37.

#### 9.4. Beziehungen zu Verkaufserfolgen

In einer Untersuchung stellte sich heraus, dass sehr signifikante Beziehungen zwischen den Faktoren UNZ, ESV, OPS und der Arbeitsleistung bestand. Die untersuchte Stichprobe bestand aus Verkäufern (N = 48). Das Außenkriterium Gehalt war in seiner Varianz hauptsächlich vom Einsatz des einzelnen abhängig. Persönlichkeits-unabhängige Einflüsse, wie Fixgehalt oder ähnliches waren nicht vorhanden. Die Korrelationen mit dem Gehaltskriterium bewegten sich zwischen .48 und .60. Diese Koeffizienten sind als ausgesprochen hoch zu bezeichnen.

Beim praktischen Einsatz einer Vorversion des DNLA-Verfahrens bei Auswahl von Versicherungsaußendienstmitarbeitern wurde davon gesprochen, dass sich die mittlere Produktivität der mit Hilfe des Tests eingestellten Personen um ca. 20 % verbesserte sowie die Fluktuation deutlich zurückging (um ca. 15-25 %). Diese Zahlen wurden nur gesprächsweise mitgeteilt, exakte Zahlen darüber liegen jedoch nicht vor.

#### 9.5. Allgemeiner Berufserfolg

Der biografische Informationsbogen (BIB) liefert (siehe STRASSER, 1975) ein Maß für persönliche Leistungsvoraussetzungen/Leistungshöhe, das methodisch völlig anders aufgebaut ist als die Faktoren des DNLA-Verfahrens. Validitäten zu einem Gesamt-Index aus dem BIB sind deswegen als wichtige Hinweise für die Validität einzelner Faktoren zu betrachten.

##### Validitäten zum BIB-Gesamtwert

UNZ -0.35

NAA -0.56

LZS 0.37

UNN -0.57

ESV 0.46

Der biografische Informationsbogen für sich gab selbst bei strengsten statistischen Kontrollen (Doppelt-Kreuz-Validierung) Validitätskoeffizienten von 0.60 bis 0.70.

#### 9.6. Beziehungen zu Schulnoten

Bei Untersuchungen mit Schulkindern ergaben sich sehr signifikante Zusammenhänge mit den Schulnoten. Die gebundenen Korrelationskoeffizienten waren teilweise höher als die Validitäten von Intelligenztests.

(Bereich 0.30 bis 0.35: Validität von Intelligenztests 0.25 - 0.30)

### 9.7. Verhaltensanalysen

In einer Untersuchung wurden die Faktoren des Sozialbereichs mit den Interaktionsverhalten in Beziehung gesetzt. Die Meßmethode des Interaktionsverhaltens war ein Test, bei dem nach Art des Rosenzweig-Picture-Frustration-Test Gesprächsergänzungen vorgenommen werden mussten. Dieser Test war ganz auf Verkaufssituationen zugeschnitten. Dabei zeigte sich sehr deutlich, dass Personen mit einem zu hohen oder zu niedrigen Selbstwertgefühl besondere Kommunikationsstile haben. Die Personen mit einem zu hohen Selbstwertgefühl behandeln andere von oben herab, man merkt dem Interaktionsverhalten deutlich an, dass sie am Kontakt und an der richtigen Behandlung des Kommunikationspartners nicht interessiert sind. Personen mit einem zu niedrigen Selbstwertgefühl tendieren dazu, sich zu unterwürfig zu geben (was für den Gesprächspartner oft peinlich ist) oder zu schnell aus einer Situation zu entfliehen. In all den geschilderten Fällen wird der Zusammenhang der Kette Selbstwertgefühl - Interaktionsverhalten - sozialer Erfolg recht deutlich.

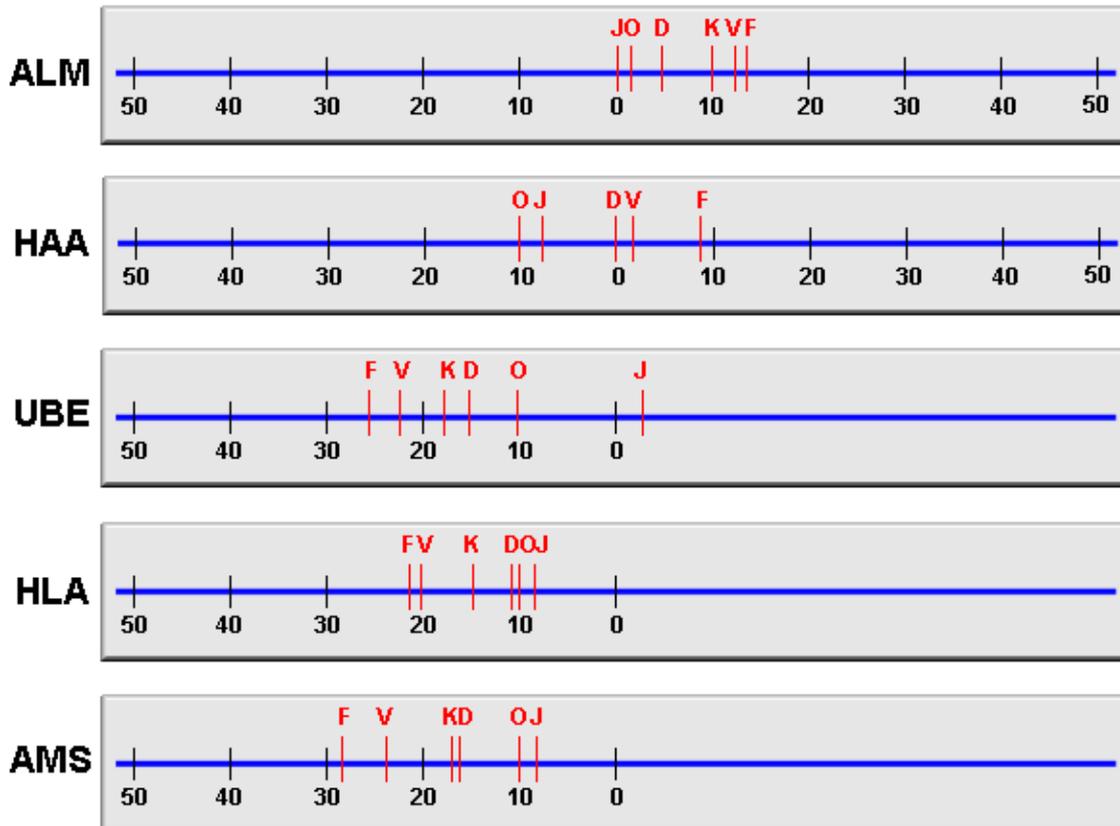
### 9.8. Extremgruppenvergleiche

Wenn der Test valide sein soll, dann muss er verschiedene Gruppen voneinander trennen können. Ferner sollten Gruppen, die sich dem allgemeinen Verständnis nach deutlich in der Leistung unterscheiden, in der "richtigen" Reihenfolge in den Testwerten erscheinen. Die Überprüfung dieser Hypothesen und die Ergebnisse, die sich dabei zeigten, sind in der folgenden Abbildung wieder-gegeben.

(Anmerkung zur Abbildung: Diese Ergebnisse wurden mit einem Vorläufer des jetzigen Bogens gefunden, so dass nicht alle Faktoren eingezeichnet sind.)

## Abbildung Faktoren des DNLA-

Position der  
einzelnen  
Gruppen auf  
den Skalen



### Stichprobenbezeichnungen:

- J = Jugendliche Verwahrloste
- O = Obdachlose (Nicht-Sesshafte)
- D = Delinquenten
- K = Kandidaten für Filialleiterpositionen  
(Lebensmitteleinzelhandel)
- V = Verkäufer
- F = Führungskräfte

### 9.9. Validität zu Vorgesetzten-Beurteilungen

Bei einer Bank wurde  $N = 58$  Auszubildende hinsichtlich der Gesamtqualifikation in eine Gruppe von "guten" und "weniger zufriedenstellenden" unterteilt. Hinsichtlich der Faktoren des DNLA-Verfahrens (die Faktoren wurden zu übergeordneten Gruppen zusammengefasst) ergaben sich folgende Ergebnisse:

1. Die WZ-Gruppe besitzt weniger Gefühl von Selbstvertrauen, Selbstwert, weniger Gefühl der Kompetenz. Dieser Faktor ist im Profil mit  $SK \ominus$  bezeichnet.
2. Die WZ-Gruppe ist angesichts von Problemen, Schwierigkeiten, Leistungssituationen ängstlicher und viel weniger stabil als die G-Gruppe. Es ist unmittelbar einsichtig, dass bei Angst und Nervosität die Qualität jeglicher Leistung (sowohl in schulischer Hinsicht wie auch im Umgang mit anderen Menschen deutlich abnimmt. Profil-Faktor Stabilität ( $ST \ominus$  )
3. Die WZ-Gruppe hat weniger als die G-Gruppe das Gefühl, selbst über Erfolg und Misserfolg bestimmen zu können. Sie hat den Eindruck, dass Erfolge und Misserfolge eher von unkontrollierbaren äußeren Einflüssen abhängen. Dies bringt geringere Erfolgserwartungen und geringere Motivation mit sich, denn Anstrengung lohnt sich ja nur dann, wenn man (wenigstens in der Mehrzahl der Fälle) den Erfolg auch beeinflussen kann. Dieser Faktor der Selbst-Bestimmung der Ergebnisse des eigenen Handelns hängt einerseits mit dem Selbstwert- und Kompetenzgefühl (SK) zustande, andererseits auch mit dem Gefühl der Eigenbestimmung (EB). In der psychologischen Literatur wird er auch „internal locus of control“ (Eigenbestimmung der Handlungs-Ergebnisse, ILC) genannt, im Gegensatz zum „external locus of control“ (Fremdbe-

stimmtheit, Schicksalsabhängigkeit, "Glücks"-abhängigkeit bei den Handlungsergebnissen) genannt.

4. Aus den bisherigen Unterschieden in den Grundfaktoren lässt sich ableiten, dass bei der WZ-Gruppe die Gesamtleistungs- bzw. Einsatzmotivation deutlich geringer sein wird als bei der G-Gruppe. Dies wird in der Item-Analyse bestätigt, wo einige MOT-Items die höchsten Trennschärfen aufweisen. Die WZ-Gruppe ist also viel weniger zu hohem Einsatz zu motivieren, was sich natürlich in Leistungen und Beurteilungen deutlich niederschlägt.
5. Die WZ-Gruppe zeigt gegenüber Einrichtungen, wie der Schule, gegenüber Lehrern und ähnlichem eine deutliche kritischere, mit mehr Unzufriedenheit und Ressentiments aufgeladene Haltung. Dies kann einerseits von mangelnden Erfolgserlebnissen herrühren, andererseits aber auch von zu überzogenen Erwartungen.
6. Die WZ-Gruppe scheint insgesamt höhere Ansprüche zu stellen als die G-Gruppe. Es ist durchaus auch möglich, dass diese Ansprüche einfach unrealistischer sind als bei der G-Gruppe.

## Validität der Faktoren

(Stichprobe Azubis)

Als Validitätskriterium, wurde die Signifikanz des Unterschieds zwischen der „G“ und „WZ“-Gruppe definiert.

Wenn der Test zwischen beiden Gruppen signifikant unterscheidet, dann kann man von einer Gültigkeit ausgehen, insbesondere da man schon eine hoch-ausgelesene Gruppe vor sich hat.

Es ist sicherlich plausibel anzunehmen dass bei der Auswahl von Bewerbern die Validität noch höher sein wird.

Faktor	Gruppe G		Gruppe WZ		Signifikanz des Unter- schieds
SP	3,67	4,62	-3,03	6,44	s.s
EB	-1,19	4,55	-6,64	4,9	s.s
GE	4,41	9,11	-2,09	11,07	s.s
MOT	4,45	10,04	-18,29	9,46	s.s
ST	25,5	14,3	-6,30	24,7	s.s
SK	10,0	14,22	-15,38	17,87	s.s

Die Ergebnisse zeigen auf allen Faktoren sehr signifikante Unterschiede.

Die Betrachtung der Varianz zeigt außerdem, dass die WZ-Gruppe deutlich heterogener ist als die G-Gruppe. Dies ist so zu interpretieren, dass die G-Gruppe insgesamt ausgeglichener ist,

wogegen bei der WZ-Gruppe eher Extremantworten durch den Prozess der "Oszillation" eines Merkmals zu erwarten sind.

10. Validitäts-Studien im europäischen Vergleich

Tabelle 1 Bedeutung der Variablen

Variable 16	Fachkompetenz	(1) hoch (3) niedrig	(2) mittel
Variable 17	Wieviel Jahre im Unternehmen	(1) ½ Jahr (3) 2-5 Jahre	(2) 1 Jahr (4) über 5 Jahre
Variable 18	Erneut einstellen?	(1) ja	(2) nein
Variable 21	Entwicklung	(1) aufsteigend (3) absteigend	(2) gleich
Variable 22	Erfolgschancen	(1) schnell (3) langfristig	(2) mittelfristig (4) keine
Variable 23	Geschlecht	(1) Mann	(2) Frau

Variable 1 – 15 : DNLA-Faktoren

Validitäts-Analyse Frankreich N = 47

Korrelationen zwischen einigen Außenkriterien und Test-Faktoren des DNLA.

1. Variable „Wieder-Einstellung“  
(Würden Sie diese Person noch einmal einstellen?  
Ja / nein)

Faktor-Name	Code	Korrelation
Auftreten	INF	0,52 ***
Status	GED	0,52 ***
Emotionale Grundhaltung	NAA	0,44 **
Selbstsicherheit	UNN	0,37 *

- \* signifikant auf dem 5 % Niveau
- \*\* signifikant auf dem 1 % Niveau
- \*\*\* signifikant auf dem 1 Promille Niveau

2. Variable „Dauer der Betriebszugehörigkeit“

½ J. / 1 J. / 2 J. / 3-4 J. / 5 und mehr Jahre

Keine signifikanten Korrelationen zu den Testwerten.

3. Variable „Fachkompetenz“

Wie hoch schätzen Sie die Fachkompetenz des Mitarbeiters ein?

Faktor-Name	Code	Korrelation
Auftreten	INF	0,34 *
Emotionale Grundhaltung	NAA	0,49 *
Selbstsicherheit	UNN	0,33

4. Variable „Geschlecht“

männlich / weiblich

keine signifikanten Korrelationen

Tabelle 4 Korrelationen DNLA-Faktoren mit Außenkriterien (Finnland)

Nr.	Faktor-Name	Code	Var. 16 Fach- Kompetenz	Var. 17 Jahre im Unternehmen	Var. 18 Erneute Einstellung	Var. 19 Andere Position	Var. 21 Ent- wicklung	Var. 22 Erfolgs- Chancen	Var. 23 Geschlecht
1	Eigenverantwortlichkeit	ILC	.51***	n.s.	.45***	-	.38**	.50**	n.s.
2	Leistungsdrang	HLA	.51***	n.s.	.29*	-	.47***	.49***	n.s.
3	Selbstvertrauen	AMS	.64***	n.s.	.50***	-	.33*	.59***	n.s.
4	Motivation	KVF	.36**	n.s.	.37**	-	.45***	.49***	n.s.
5	Selbstbewusstsein	ASB	.35**	n.s.	n.s.	-	n.s.	n.s.	n.s.
6	Kontaktfähigkeit	DIS	.31*	n.s.	n.s.	-	n.s.	n.s.	n.s.
7	Auftreten	INF	.31**	n.s.	n.s.	-	n.s.	.30*	.28*
8	Einsatzfreude	ESV	.33*	n.s.	.37**	-	.29*	.47***	n.s.
9	Anspruchsniveau	LZS	.29*	n.s.	.36**	-	.38**	.56***	n.s.
10	Status/Prestige	GED	.28*	n.s.	n.s.	-	n.s.	n.s.	n.s.
11	Gelassenheit	GEB	.27*	n.s.	n.s.	-	n.s.	n.s.	.33*
12	Kritikstabilität	KST	.29*	n.s.	.26*	-	.27*	.37*	.27*
13	Misserfolgstoleranz	MET	.29*	n.s.	.26*	-	.27*	.36**	.29*
14	Emotionale Grundhaltung	NAA	.37*	n.s.	.44***	-	.46***	.53***	n.s.
15	Selbstsicherheit	UNN	.47***	n.s.	.27*	-	.30*	.42***	n.s.

\*           signifikant auf dem 5 % Niveau  
 \*\*          signifikant auf dem 1 % Niveau  
 \*\*\*        signifikant auf dem 1 Promille Niveau

## INTERPRETATION

Stichprobe Finnland:

Die vom Vorgesetzten beurteilte Fach-Kompetenz (Var 16) korreliert signifikant mit allen DNLA-Faktoren. Offenbar ist dieses Außenkriterium ein allgemeiner Wohlwollens-Faktor seitens des Vorgesetzten.

Auf dem 1-Promille-Niveau korrelieren die Faktoren ILC, HLA, AMS und UNN. Dies könnte so interpretiert werden, als würde das Wohlwollen des Vorgesetzten hauptsächlich mit einigen Faktoren der emotionalen Stabilität verknüpft sein. Nach den vorliegenden Daten schätzen Vorgesetzte also Personen insgesamt höher ein, die als „weniger schwierig“ zu bezeichnen sind.

Mit der Günstigkeit der Entwicklung (Var. 21) korrelieren vor allem die Faktoren HLA, KVF und NAA; sowie ILC und LZS. Das bedeutet, dass hier offenbar Faktoren der Leistungsmotivation eine größere Rolle spielen.

Bezüglich der Einschätzung der Erfolgs-Chancen (Var. 22) heben sich offenbar noch deutlicher die Faktoren der Leistungsmotivation hervor. Hochsignifikante Korrelationen sind vor allem bei den Faktoren ILC, KVF, ESV, LZS zu finden. Aber auch die Stabilitäts-Faktoren weisen sehr hohe Korrelationen auf: HLY, AMS, NAA und UNN.

Die Plausibilität der Daten wird dokumentiert durch die Tatsache, dass die Dauer der Zugehörigkeit zum Unternehmen durchweg keine signifikanten Korrelationen aufweist.

Zusammenhänge zum Geschlecht der Testpersonen sind ebenfalls sehr gering. Frauen weisen allerdings signifikant niedrigere Werte in den Faktoren INF, GEB, KST, MET auf. Offenbar fühlen Frauen sich unterlegener, weniger gelassen, weniger stabil gegenüber Kritik und Misserfolgen. Dies ist offenbar auch im Bereich hoher Plausibilität.

## Zusammenfassung

Diese Daten zeigen für die einzelnen Faktoren des Tests eine genügend hohe Validität, um einen Einsatz zur Entscheidungshilfe bei Personalauswahl voll zu rechtfertigen.

## Validitäten Italien

Wie auch bei den anderen Validitätsstudien werden signifikante Korrelationen zwischen einigen Testfaktoren und dem Außenkriterium (Erfolg in einer neuen Position) gefunden.

Tabelle: Validitätsstudie Italien

Nr.	Faktor-Name	Code	Var. 22 Erfolgschancen
1	Eigenverantwortlichkeit	ILC	-.15
2	Leistungsdrang	HLA	-.24
3	Selbstvertrauen	AMS	-.27
4	Motivation	KVF	-.03
5	Selbstbewusstsein	ASB	-.01
6	Kontaktfähigkeit	DIS	-.19
7	Auftreten	INF	-.39*
8	Einsatzfreude	ESV	-.18
9	Anspruchsniveau	LZS	-.36*
10	Status / Prestige	GED	-.34*
11	Gelassenheit	GEB	-.39*
12	Kritikstabilität	KST	-.25
13	Misserfolgstoleranz	MET	-.05
14	Emotionale Grundhaltung	NAA	-.37*
15	Selbstsicherheit	UNN	-.24

\* = signifikant 1 %

DNLA korreliert besonders mit geschätzter Erfolgsmöglichkeit in Positionen.

INF, LZS, GED, GEB, NAA

Stichprobe Italien, Validitäten zur Vorgesetztenbeurteilung  
N = 49, August 1989

Nr.	Faktor-Name	Code	Fach-Kompetenz	Erfolg in neuer Position	Geschlecht	Jahre in Unternehmen
1	Einsatzfreude	ESV	-	-	-.45***	-
2	Status/Prestige	GED	-	-	-	-
3	Eigenverantwortlichkeit	ILC	-	-	-	-
4	Anspruchsniveau	LZS	-	-.32*	-	-
5	Selbstvertrauen	AMS	-	-.26*	-.31*	-
6	Leistungsdrang	HLA	-.39*	-.28*	-.28*	-
7	Motivation	KVF	-	-	-	-
8	Gelassenheit	GEB	-	-.45***	-.33*	-
9	Kritikstabilität	KST	-.28*	-.41**	-	-
10	Misserfolgstoleranz	MET	-	-	-	-
11	Emotionale Grundhalt.	NAA	-	-	-	-
12	Selbstsicherheit	UNN	-	-.39**	-	-
13	Zufriedenheit	UNZ	-	-.39**	-	-
14	Selbstbewusstsein	ASB	-	-	-	-
15	Kontaktfähigkeit	DIS	-.27*	-	-	-
16	Auftreten	INF	-	-	-	-

- \* = signifikant auf 5 % Niveau
- \*\* = signifikant auf 1 % Niveau
- \*\*\* = signifikant auf 0,1 % Niveau

Besondere Validität weisen offenbar die Faktoren Anspruchsniveau, Selbstvertrauen, Leistungsdrang, Gelassenheit, Selbstsicherheit und Zufriedenheit auf. Die Plausibilität der Daten zeigt sich in den nicht-signifikanten Korrelationen zu der Unternehmens-Zugehörigkeits-Dauer.

Die Analyse der Fremdeinschätzung zeigt auf allen Faktoren hochsignifikante Korrelationen. Das bedeutet, dass (wie bei Einschätzungen üblich) kaum Differenzierungen zwischen den Begriffen gemacht werden, sondern mehr oder weniger alles von einem großen Halo-Effekt überstrahlt wird. Als brauchbare Einschätzung der wirklichen Ausprägung eines Merkmals kann also das Vorgesetztenurteil nicht dienen, es ist eine allgemeine, eher undifferenzierte Qualifikations-Aussage.

Tabelle: Korrelationen zwischen der Einschätzung der Faktoren durch die Vorgesetzten und dem Kriterium „Erfolg in einer neuen Position“

Nr.	Faktor-Name	Code	Var. 22 Erfolgschancen
1	Eigenverantwortlichkeit	ILC	-.47**
2	Leistungsdrang	HLA	-.42*
3	Selbstvertrauen	AMS	-.57**
4	Motivation	KVF	-.50**
5	Selbstbewusstsein	ASB	-.48**
6	Kontaktfähigkeit	DIS	-.40*
7	Auftreten	INF	-.58**
8	Einsatzfreude	ESV	-.52**
9	Anspruchsniveau	LZS	-.55**
10	Status / Prestige	GED	-.34*
11	Gelassenheit	GEB	-.35*
12	Kritikstabilität	KST	-.46**
13	Misserfolgstoleranz	MET	-.50**
14	Emotionale Grundhaltung	NAA	-.37*
15	Selbstsicherheit	UNN	-.55**

\* = signifikant auf dem 1 % Niveau  
 \*\* = signifikant auf dem 0,1 % Niveau

## 11. Akzeptanz

Das valideste Messinstrument hätte keinen praktischen Nutzen, wenn keine Akzeptanz durch die Benutzer oder durch die Kandidaten vorhanden wäre.

Akzeptanz besteht im wesentlichen in den folgenden Bedingungen:

### 1. Fairness der Daten-Erhebung

Klarheit der Spielregeln und Prozeduren (keine "Fallen"; kein unnötiger Stress; keine Informationen, über die der Kandidat nichts weiß.)

### 2. Kein Gesichtsverlust

Jeder Kandidat muss die Möglichkeit haben, eine gute Beratung zu bekommen, ohne dass dabei (im Extremfall) ein "Gesichtsverlust" entsteht.

### 3. Feedback

Jeder Kandidat, der sich einem Auswahlverfahren unterzieht, sollte alle Ergebnisse des Verfahrens bekommen. Voll gewährleistet.

### 4. Ethische Aspekte

Keine Auswertungen, die den Privatbereich betreffen oder Aussagen über Verhaltensweisen/Eigenschaften, die nichts mit dem Beruf zu tun haben. ("Unkontrollierte" Methoden wie Graphologie, Interview, etc. gehen hier oft viel zu weit!)

### 5. Akzeptanz

#### Datenschutz

Keine personenbezogenen Daten im elektronischen Datenträger gespeichert.

### "Image" der Prozedur und des Gutachtens

Ein sehr wichtiger, bisher viel zu wenig beachteter Punkt für die Akzeptanz der Auswahlmethodik insgesamt und damit für das Gelingen der Auswahl ist das "Image" des Auswahlprozesses. Durch Interviews wurde herausgefunden, dass vor allem qualifizierte Bewerber sich für dasjenige Unternehmen entscheiden, bei dem sie sich am besten behandelt fühlen. Bewerber erwarten eine qualifizierte Potentialanalyse und Beratung darüber, ob das eigene Fähigkeitenprofil mit den jeweiligen Anforderungen übereinstimmt. Ein "Massentest" mit einem unqualifizierten, wenig erklärten Testprofil als Aussage (bzw. kein Feedback über das Abschneiden) ist deshalb wenig geeignet, ein gutes Image zu hinterlassen.

Deshalb wurde beim DNLA besonderer Wert auf eine individuelle Durchführbarkeit des Tests sowie ein qualitativ sehr hochwertiges Gutachten mit sehr vielen Erklärungen gelegt. Die bisherigen Beobachtungen zeigen, dass DNLA sowohl vom Testvorgang als auch vom Gutachten her als ausgesprochen positiv eingestuft wird.

Das DNLA-Gutachten als Grundlage für Einarbeitung  
Führung und Entwicklung der neuen Mitarbeiter

Einige Faktoren der Kompetenz sind veränderbar, andere nicht.

Wenig veränderbar sind z.B. Faktoren wie allgemeine Intelligenz oder Energie.

Andere Faktoren sind entwicklungsfähig.

Deshalb ist es wichtig, dass eine Potential-Analyse Hinweise für Führung bzw. Weiterentwicklung enthält.

Gerade dieser Entwicklungsaspekt wird sowohl von den Bewerbern wie auch den Führungskräften mit am meisten geschätzt.

Das DNLA-Gutachten ist eine hervorragende Basis für das Einarbeitungsgespräch und die gezielte Entwicklung für die Mitarbeiter. Es kann „blinde Flecke“, Bereiche bisher nicht bekannter "Erfolgs-Bremsen" aufdecken, sowie mehr Sicherheit über die eigenen Stärken geben.

## ANHANG 1

Einige Item-Beispiele aus anderen Tests, die

- \* weit weg von beruflichen Angelegenheiten sind
- \* unserer Meinung nach zu weit in den Privatbereich hineingehen

(Aus dem 16 PF)

Wilde Tiere lösen in mir ein gewisses Unbehagen aus, selbst wenn sie in stabilen Käfigen sind.

- (a) ja, stimmt      (b) unsicher   (c) nein, stimmt nicht

Leichte klassische Musik ziehe ich Schlagern vor.

- (a) stimmt    (b) unsicher   (c) stimmt nicht

Wenn ich sähe, wie zwei Kinder meiner Nachbarn sich streiten,

- (a)    ließe ich sie ihren Streit unter sich ausmachen  
(b)    unsicher  
(c)    redete ich ihnen vernünftig zu

An einem freien Abend würde ich lieber

- (a)    ein gutes Buch lesen  
(b)    unsicher  
(c)    mit Freunden an einem Hobby arbeiten

Ich würde lieber jemanden heiraten, der

- (a)    allgemein bewundert wird  
(b)    unsicher  
(c)    künstlerisch oder literarisch begabt ist

Ich ertappe mich dabei, wie ich Dinge ohne besonderen Anlaß zähle.

- (a) oft            (b) gelegentlich      (c) nie

Bekanntem gegenüber ziehe ich es vor,

- (a)    mich auf sachliche unpersönliche Gespräche zu beschränken  
(b)    dazwischen  
(c)    mich über andere Menschen und deren Meinungen zu unterhalten

In meinem Privatleben erreiche ich fast immer die Ziele, die ich mir setze.

- (a) stimmt    (b) unsicher   (c) stimmt nicht

Viele einfache Menschen wären entsetzt, würden sie meine inneren persönlichen Ansichten kennen.

- (a) ja            (b) unsicher      (c) nein

Ich freue mich immer, wenn ich an einer großen gesellschaftlichen Veranstaltung teilnehme, sei es eine Party, ein Tanzabend oder ein öffentliches Treffen.

- (a) ja            (b) dazwischen            (c) nein

Die meisten Leute "spinnen" ein bißchen, obwohl sie es nicht gerne zugeben wollen.

- (a) ja            (b) dazwischen            (c) nein

Ich mag einen Freund (meinen Geschlechts), der

- (a) sich ernsthafte Gedanken über seine Lebenseinstellung macht  
(b) dazwischen  
(c) tüchtig und praktisch ausgerichtet ist

Auf intellektuellem Gebiet sind (waren) meine Eltern

- (a) durchschnittlich            (b) unsicher            (c) über dem Durchschnitt

Ich empfinde ein starkes Verlangen nach jemandem, auf den ich mich verlassen kann, wenn ich niedergeschlagen bin.

- (a) ja            (b) dazwischen            (c) nein

Wenn ich in einen Spiegel sehe, wundere ich mich manchmal über die Bedeutung von rechts und links.

- (a) stimmt    (b) unsicher    (c) stimmt nicht

Als Jugendlicher habe ich Sport getrieben

- (a) gelegentlich    (b) ziemlich oft    (c) sehr häufig

Auf der Straße würde ich eher stehenbleiben, um einem Künstler beim Malen zuzuschauen, als um Leuten zuzuhören, die sich streiten.

- (a) stimmt            (b) unsicher            (c) stimmt nicht

Ich bewundere meine Eltern in allen wesentlichen Funktionen.

- (a) ja            (b) unsicher            (c) nein

Ich gehe abends gern aus.

- (a) öfter als einmal in der Woche (mehr als der Durchschnitt)  
(b) ungefähr einmal in der Woche (durchschnittlich)  
(c) weniger als einmal in der Woche (weniger als der Durchschnitt)

Ich glaube, dass ein großes Maß an Freiheit wichtiger ist als gute Manieren und Achtung vor dem Gesetz.

- (a) stimmt (b) unsicher (c) stimmt nicht

Ich würde lieber in einer Stadt leben,

- (a) die geschäftlich äußerst rege und im Aufstieg ist, mit einer etwas rauen Atmosphäre  
(b) unsicher  
(c) die künstlerisch angelegt, aber verhältnismäßig arm ist

Ich neige dazu, viele unwichtige, belanglose Dinge zu vergessen, wie Namen von Straßen oder Geschäften in der Stadt.

- (a) ja (b) dazwischen (c) nein

Ich esse mit Genuss, aber nicht immer so achtsam und korrekt, wie manche Leute es tun.

- (a) stimmt (b) unsicher (c) stimmt nicht

Wenn ich als Jugendlicher anderer Meinung war als meine Eltern,

- (a) blieb ich gewöhnlich bei meiner Meinung  
(b) dazwischen  
(c) erkannte ich gewöhnlich ihre Überlegenheit an

Ich würde lieber jemanden heiraten,

- (a) der dafür sorgt, dass sich die Familie in ihrem eigenen Interessenkreis beschäftigt  
(b) dazwischen  
(c) der die Familie am gesellschaftlichen Leben der Umgebung teilnehmen lässt

Ordinäre Redensarten stoßen mich ab, auch dann, wenn sie nicht in einer aus Männern und Frauen zusammengesetzten Gesellschaft gebraucht werden.

- (a) ja (b) dazwischen (c) nein

Ich wäre sehr ungern an einem Ort, wo nicht eine Menge Leute sind, mit denen ich sprechen könnte.

- (a) stimmt (b) unsicher (c) stimmt nicht

Meine Zurückhaltung steht mir immer im Wege, wenn ich mit einem anziehenden, aber mit fremden Menschen (des anderen Geschlechts) sprechen möchte.

- (a) ja (b) dazwischen (c) nein

Diskussionen über Kunst, Religion oder Politik erregen mich selten so sehr, daß ich Höflichkeit und Benehmen vergesse.

- (a) stimmt    (b) unsicher    (c) stimmt nicht

Ich würde Maßnahmen begrüßen, die dafür Sorge tragen,

- (a) dass mehr pflanzliche Nahrung gegessen wird, damit nicht so viele Tiere getötet werden müssen  
(b) unsicher  
(c) dass man bessere Schädlingsbekämpfungsmittel entwickelt, um die Tiere zu töten, die die Ernte des Bauern vernichten (wie Eichhörnchen, Kaninchen und manche Vogelarten)

Im Alter von ungefähr 16 Jahren ging ich mit Bekannten des anderen Geschlechts aus.

- (a) sehr häufig  
(b) so oft wie andere Jugendliche  
(c) sehr selten

Die Auffassung, dass Krankheiten ebenso viele seelische wie körperliche Ursachen haben, ist ziemlich Übertrieben.

- (a) ja    (b) dazwischen    (c) nein

Ich habe mindestens soviel Freunde des anderen Geschlechts wie des eigenen.

- (a) ja    (b) dazwischen    (c) nein

Auch in einem wichtigen Spiel geht es mehr darum, mich dabei zu unterhalten, als es zu gewinnen.

- (a) immer    (b) im allgemeinen    (c) gelegentlich

Ich wäre gerne ein Zeitungskritiker für Dramen, Konzerte, Opern usw.

- (a) ja    (b) unsicher    (c) nein

Ich glaube, in der heutigen Welt ist es viel wichtiger, folgendes zu lösen:

- (a) die politischen Spannungen  
(b) unsicher  
(c) die Frage der sittlichen Lebensführung

Als Kind fürchtete ich mich im Dunkeln.

- (a) oft    (b) manchmal    (c) nie