



universität
wien

Diplomarbeit

Titel der Diplomarbeit

Evaluation zweier englischer Sprachtests

Verfasserin

Petra Laa

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Schwechat, im September 2008

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuer: a.o. Univ.-Prof. Dr. Mag. Martin E. Arendasy

DANKSAGUNG

An dieser Stelle möchte ich mich recht herzlich bei all jenen bedanken, die mich nicht nur während meines Studiums, sondern auch im Laufe meines gesamten Lebens freundlich unterstützt haben.

Zu allererst möchte ich meiner Mutter, Irene Laa, ein großes Dankeschön aussprechen, die mir nicht nur finanziell all die Jahre unter die Arme gegriffen hat, sondern mir auch jeder Zeit mit Rat und Tat zur Seite stand.

Bei meinem Freund, DI Stefan Stracke, möchte ich mich ebenfalls für seine enorme Geduld mit mir bedanken. Vor allem gegen Ende des Studiums lagen die Nerven oft blank, doch Dank seiner kompetenten Unterstützung bei der Datenverarbeitung konnten auch noch so schwierige Probleme gelöst werden.

Ein großer Dank gebührt auch meinen Freundinnen Andrea Fleischmann, Julia Strobl und Birgit Oberortner, die jeder Zeit ein offenes Ohr für meine Probleme hatten.

Last but not least möchte ich auch noch meinem Vater, Ing. Manfred Laa, meinen Dank aussprechen, der durch seine unglaublich positive Lebenseinstellung immer optimistische Stimmung verbreitet hat. Auch wenn er das Ende meines Studiums leider nicht mehr miterleben konnte, bleibt die Erinnerung an einen wundervollen Menschen und ich würde mir wünschen, nur einen Funken dieser positiven Lebenseinstellung übernehmen zu können.

INHALTSVERZEICHNIS

| | |
|---|-----------|
| I. EINLEITUNG..... | 1 |
| II. THEORETISCHER TEIL | 3 |
| 1. Intelligenzstrukturmodelle | 3 |
| 1.1. Die Zwei-Faktoretheorie bzw. die Generalfaktoretheorie von Spearman..... | 4 |
| 1.2. Die Primärfaktoretheorie von Thurstone..... | 5 |
| 1.3. Das morphologische Modell von Guilford..... | 7 |
| 1.4. Die Theorie der fluiden und kristallinen Intelligenz von Cattell | 8 |
| 1.5. Das Berliner Intelligenzstrukturmodell (BIS) von Jäger..... | 10 |
| 1.6. Die Three-Stratum-Theory von Carroll..... | 12 |
| 1.7. Die Theorie der multiplen Intelligenzen von Gardner | 14 |
| 2. Intelligenzdiagnostik und Diagnostik verbaler Fähigkeiten | 17 |
| 2.1. Die Anfänge der Intelligenzdiagnostik | 17 |
| 2.2. Dimensionen sprachlicher Fähigkeiten | 21 |
| 2.2.1. Analogien (ANA)..... | 22 |
| 2.2.2. Antonyme (ANT) | 23 |
| 2.2.3. Assoziationen (ASS) | 23 |
| 2.2.4. Klassifikation (KLA)..... | 24 |
| 2.2.5. Komplexe sprachliche Inhalte (KSI)..... | 25 |
| 2.2.6. Merkfähigkeit (MEK) | 26 |
| 2.2.7. Rechtschreibkenntnisse (REK) | 27 |
| 2.2.8. Selektion (SEL)..... | 27 |
| 2.2.9. Unmöglichkeiten (UNM) | 28 |
| 2.2.10. Vorstellungsvermögen (VOR) | 29 |
| 2.2.11. Wiedererkennen (WEK)..... | 30 |
| 2.2.12. Wortverständnis (WV) | 30 |
| 2.2.13. Synonyme (SYN) | 31 |
| 2.2.14. Flüssigkeit (FLU) | 34 |
| 3. Mentales Lexikon | 42 |
| 3.1. Methoden zur Erforschung der Struktur des mentalen Lexikons..... | 44 |
| 3.1.1. Sprechfehler | 44 |
| 3.1.2. Wortfindungsstörung bei hirngesunden Sprechern | 45 |
| 3.1.3. Wortfindungsstörung bei amnestischer Aphasie..... | 45 |
| 3.1.4. Chronometrische Studien | 46 |
| 3.2. Die Bedeutung von Wörtern | 47 |
| 3.2.1. Die Checklistentheorie | 47 |
| 3.2.2. Prototypentheorie | 48 |
| 3.3. Die Verbindung zwischen den Wörtern | 48 |
| 3.4. Zugriff auf das mentale Lexikon..... | 50 |
| 3.4.1. Die Wortproduktion | 51 |
| 3.4.2. Das Erkennen akustischer Wörter | 55 |
| 3.5. Die Gesamtstruktur des mentalen Lexikons..... | 57 |

| | |
|--|------------|
| 4. Der kindliche Spracherwerb | 60 |
| 4.1. Erkennen der Wortbedeutung bei Kindern..... | 60 |
| 4.1.1. Etikettieren | 60 |
| 4.1.2. Sortieren | 62 |
| 4.1.3. Aufbau des Netzwerkes..... | 63 |
| 4.2. Die Phasen des Spracherwerbs..... | 63 |
| 4.2.1. Die ersten 50 Wörter | 63 |
| 4.2.2. Der schnelle Erwerb des Wortschatzes | 64 |
| 5. Einflussfaktoren auf intellektuelle bzw. sprachliche Fähigkeiten | 67 |
| 5.1. Alter..... | 67 |
| 5.2. Geschlecht | 68 |
| 5.3. Bildung | 69 |
| III. EMPIRISCHER TEIL | 71 |
| 6. Testmaterial und Durchführung der Testung | 71 |
| 7. Synonyme Fragebogen engl (Version 1)..... | 74 |
| 7.1. Beschreibung der Stichprobe | 74 |
| 7.1.1. Auswahl der gültigen Testpersonen | 74 |
| 7.1.2. Bearbeitungsdauer | 74 |
| 7.1.3. Alter..... | 75 |
| 7.1.4. Geschlecht | 76 |
| 7.1.5. Bildung | 77 |
| 7.2. Reliabilität und Trennschärfen (N = 73) | 78 |
| 7.3. Verteilung des Rohscores..... | 79 |
| 7.4. Überprüfung auf Rasch-Homogenität | 81 |
| 7.4.1. Teilungskriterium Geschlecht | 82 |
| 7.4.2. Teilungskriterium Alter | 84 |
| 7.4.3. Teilungskriterium Rohscore | 85 |
| 7.4.4. Personenparameter im Synonyme-Test..... | 87 |
| 8. Word Fluency Test „WF engl“ | 89 |
| 8.1. Beschreibung der Stichprobe | 89 |
| 8.1.1. Auswahl der gültigen Testpersonen | 89 |
| 8.1.2. Bearbeitungsdauer | 89 |
| 8.1.3. Alter..... | 90 |
| 8.1.4. Geschlecht | 92 |
| 8.1.5. Bildung | 92 |
| 8.2. Reliabilität und Trennschärfen (N = 37) | 93 |
| 8.3. Verteilung des Rohscores..... | 94 |
| 8.4. Überprüfung auf Rasch- Homogenität | 95 |
| 8.4.1. Teilungskriterium Geschlecht | 97 |
| 8.4.2. Teilungskriterium Alter | 98 |
| 8.4.3. Teilungskriterium Rohscore | 100 |
| 8.4.4. Personenparameter im Wortflüssigkeitstest | 101 |
| 9. Hypothesentestung | 103 |
| 9.1. Zusammenhang der Leistungen in beiden Tests | 103 |
| 9.2. Varianzanalyse - Synonyme Fragebogen engl (Version 1)..... | 105 |

| | |
|--|------------|
| 9.3. Varianzanalyse - Word Fluency Test „WF engl“ | 108 |
| 10. Interpretation der Ergebnisse und Diskussion..... | 111 |
| 11. Zusammenfassung..... | 114 |
| LITERATURVERZEICHNIS | 116 |

ABBILDUNGSVERZEICHNIS

| | |
|--|-----|
| Abbildung 1: „Strukturmodell der Intelligenz nach Guilford“ | 8 |
| Abbildung 2: „Berliner Intelligenzstrukturmodell nach Jäger“ | 10 |
| Abbildung 3: „Three-Stratum-Theorie nach Carroll“ | 14 |
| Abbildung 4: „Dimensionen verbal-intelligenter Fähigkeiten nach Acker 2001“ | 22 |
| Abbildung 5: „DTD Subtest 1 (Passiver Wortschatz)“ | 27 |
| Abbildung 6: „Beispiel aus dem Verbalen Kurz-Intelligenztest“ | 31 |
| Abbildung 7: „Beispielitem aus dem LEWITE“ | 34 |
| Abbildung 8: „Ein Trittsteinmodell mit Engpaß“ | 52 |
| Abbildung 9: „Ein Wasserfallmodell“ | 53 |
| Abbildung 10: „Ein Interaktives Aktivierungs-Modell“ | 55 |
| Abbildung 11: „Ein ‚Luftbild‘ der Verbindungen zu anderen Komponenten“ | 59 |
| Abbildung 12: Bearbeitungsdauer im Synonyme-Test in Minuten | 75 |
| Abbildung 13: Verteilung des Lebensalters im Synonyme-Test | 76 |
| Abbildung 14: Verteilung der höchsten abgeschlossenen Ausbildung im Synonyme-Test ... | 78 |
| Abbildung 15: Verteilung des Rohscores im Synonyme-Test | 80 |
| Abbildung 16: Grafische Modellprüfung im Synonyme-Test internes Teilungskriterium.... | 87 |
| Abbildung 17: Bearbeitungsdauer im Wortflüssigkeitstest in Minuten..... | 90 |
| Abbildung 18: Verteilung des Lebensalters im Wortflüssigkeitstest..... | 91 |
| Abbildung 19: Verteilung der höchsten abgeschlossenen Ausbildung im WF-Test | 93 |
| Abbildung 20: Verteilung des Rohscores im Wortflüssigkeitstest | 95 |
| Abbildung 21: Grafische Modellprüfung im WF-Test Teilungskriterium Geschlecht..... | 98 |
| Abbildung 22: Grafische Modellprüfung im WF-Test Teilungskriterium Rohscore | 101 |
| Abbildung 23: Mittelwerte im Synonyme-Test in Abhängigkeit von Geschlecht und Alterskategorie | 107 |
| Abbildung 24: Mittelwerte im Wortflüssigkeitstest in Abhängigkeit von Geschlecht und Alterskategorie | 110 |

TABELLENVERZEICHNIS

| | |
|--|-----|
| Tabelle 1: Deskriptivstatistische Kennzahlen zur Testbearbeitungsdauer im SYN-Test..... | 75 |
| Tabelle 2: Deskriptivstatistische Kennzahlen zum Alter der Testpersonen im SYN-Test..... | 75 |
| Tabelle 3: Verteilung und Anteilswerte des Lebensalters im Synonyme-Test | 76 |
| Tabelle 4: Verteilung und Anteilswerte des Geschlechts im Synonyme-Test | 77 |
| Tabelle 5: Verteilung und Anteilswerte der höchsten abgeschlossenen Ausbildung im Synonyme-Test..... | 77 |
| Tabelle 6: Itemkennwerte im Synonyme-Test..... | 79 |
| Tabelle 7: Deskriptivstatistischen Kennzahlen zum Rohwert im Synonyme-Test | 80 |
| Tabelle 8: Schwierigkeitsparameter und Standardfehler der einzelnen Items im SYN-Test.. | 81 |
| Tabelle 9: Likelihood-Quotiententest mittels χ^2 nach Andersen im Synonyme-Test (Teilungskriterium Geschlecht)..... | 83 |
| Tabelle 10: Itemparameter der Teilstichproben („weiblich“ und „männlich“) mit Signifikanzbeurteilung nach Wald im Synonyme-Test | 83 |
| Tabelle 11: Likelihood-Quotiententest mittels χ^2 nach Andersen im Synonyme-Test (Teilungskriterium Alter) | 84 |
| Tabelle 12: Itemparameter der Teilstichproben („jung“ und „alt“) mit Signifikanzbeurteilung nach Wald im Synonyme-Test | 85 |
| Tabelle 13: Likelihood-Quotiententest mittels χ^2 nach Andersen im Synonyme-Test (Teilungskriterium Rohscore) | 86 |
| Tabelle 14: Rohwerte und die entsprechenden Fähigkeitsparameter (Θ) im Synonyme-Test | 88 |
| Tabelle 15: Deskriptivstatistische Kennzahlen zur Testbearbeitungsdauer im Wortflüssigkeitstest..... | 90 |
| Tabelle 16: Deskriptivstatistische Kennzahlen zum Alter der Testpersonen im Wortflüssigkeitstest..... | 91 |
| Tabelle 17: Verteilung und Anteilswerte des Lebensalters im Wortflüssigkeitstest | 91 |
| Tabelle 18: Verteilung und Anteilswerte des Geschlechts im Wortflüssigkeitstest | 92 |
| Tabelle 19: Verteilung und Anteilswerte der höchsten abgeschlossenen Ausbildung im Wortflüssigkeitstest..... | 93 |
| Tabelle 20: Itemkennwerte im Wortflüssigkeitstest..... | 94 |
| Tabelle 21: Deskriptivstatistische Kennzahlen zum Rohwert im Wortflüssigkeitstest..... | 95 |
| Tabelle 22: Schwierigkeitsparameter und Standardfehler der einzelnen Items im Wortflüssigkeitstest (Gesamtstichprobe) | 96 |
| Tabelle 23: Likelihood-Quotiententest mittels χ^2 nach Andersen im Wortflüssigkeitstest (Teilungskriterium Geschlecht)..... | 97 |
| Tabelle 24: Likelihood-Quotiententest mittels χ^2 nach Andersen im Wortflüssigkeitstest (Teilungskriterium Alter) | 99 |
| Tabelle 25: Itemparameter der Teilstichproben („jung“ und „alt“) mit Signifikanzbeurteilung nach Wald im Wortflüssigkeitstest | 99 |
| Tabelle 26: Likelihood-Quotiententest mittels χ^2 nach Andersen im Wortflüssigkeitstest (Teilungskriterium Rohscore) | 100 |
| Tabelle 27: Rohwerte und die entsprechenden Fähigkeitsparameter (Θ) im Wortflüssigkeitstest..... | 102 |
| Tabelle 28: Beurteilung der Normalverteilung der Daten im Synonyme-Test (SYN) und Wortflüssigkeitstest (WF) | 104 |
| Tabelle 29: Deskriptivstatistische Kennwerte zum Synonyme-Test (SYN) und Wortflüssigkeitstest (WF) | 104 |

| | |
|--|-----|
| Tabelle 30: Beurteilung der Normalverteilung der Daten in den vier Faktorstufenkombinationen im Synonyme-Test | 105 |
| Tabelle 31: Deskriptivstatistik zum Synonyme-Test in Abhängigkeit von Alter und Geschlecht | 106 |
| Tabelle 32: Tafel der Varianzanalyse im Synonyme-Test | 106 |
| Tabelle 33: Beurteilung der Normalverteilung der Daten in den vier Faktorstufenkombinationen im Wortflüssigkeitstest | 108 |
| Tabelle 34: Deskriptivstatistik zum Wortflüssigkeitstest in Abhängigkeit von Alter und Geschlecht | 109 |
| Tabelle 35: Tafel der Varianzanalyse im Wortflüssigkeitstest..... | 109 |

I. EINLEITUNG

Ausgangspunkt der hier vorliegenden Diplomarbeit ist die Evaluierung zweier englischer Sprachtests, welche von Dr. Arendasy entwickelt wurden. Der erste Test (Synonyme Fragebogen engl-Version 1) befasst sich mit der Messung des Wortschatzes anhand des Erkennens von synonymen Begriffen, während das zweite Verfahren (WF engl) das Konstrukt der Wortflüssigkeit mittels Buchstabenvertauschung zu erfassen trachtet. Im ersten Teil dieser Arbeit wird der theoretische Hintergrund der Intelligenz und der sprachlichen Fähigkeiten näher betrachtet, während es sich bei dem zweiten Teil explizit um die Evaluation der beiden Tests handelt.

Nachdem die sprachlichen Fähigkeiten einen wesentlichen Teilbereich der Intelligenz darstellen und im Laufe der Geschichte viele unterschiedliche Annahmen zur Struktur der Intelligenz entwickelt wurden, habe ich versucht im ersten Teil meiner Arbeit einen Überblick über die bedeutendsten Modelle der Intelligenzstruktur zu geben. Im Anschluss daran sollen kurz die Anfänge der Intelligenzdiagnostik erläutert werden. Aufgrund der facettenreichen Möglichkeiten die verbale Intelligenz einer Person zu testen, war es mir ein Anliegen sämtliche Dimensionen sprachlicher Fähigkeiten anzuführen und zum besseren Verständnis mit Beispielaufgaben näher zu erklären. Den beiden Dimensionen „Synonyme“ und „Flüssigkeit“ wurden aufgrund der zu evaluierenden Tests besondere Aufmerksamkeit geschenkt. Um sich eine Vorstellung davon machen zu können, was während des Sprechens in unseren Köpfen vorgeht, befasst sich Kapitel drei mit der Organisation des Mentalen Lexikons, sowie weiterer Ausführungen zur Erfassung der Wortbedeutung und Theorien über den Zugriff auf den menschlichen Wortspeicher bei der Wortproduktion und Worterkennung. Der Spracherwerb in der frühen Kindheit soll in Kapitel vier näher beschreiben, wie ein

solches Netzwerk möglicherweise aufgebaut werden kann. Schließlich beschäftigt sich Kapitel fünf noch mit den Einflussfaktoren auf intellektuelle und sprachliche Fähigkeiten.

Wie schon erwähnt sollen nun im zweiten Teil dieser Arbeit beide Tests näher betrachtet werden. Zunächst werden beide Stichproben deskriptivstatistisch überprüft und im Anschluss daran folgen Analysen zu Reliabilitäten, Trennschärfen und Rasch-Konformität. Weiters erfolgen inferenzstatistische Berechnungen zu den Einflussfaktoren Alter, Bildung und Geschlecht sowie zum Zusammenhang zwischen den beiden Tests.

Eine abschließende Interpretation und Diskussion sollen die vorliegenden Ergebnisse kritisch hinterfragen.

II. THEORETISCHER TEIL

1. Intelligenzstrukturmodelle

Die Intelligenz ist eines der am häufigsten erforschten Eigenschaften des Menschen.

Aufgrund der Tatsache, dass dieses Konstrukt nicht direkt beobachtbar ist, kann man lediglich versuchen über das Verhalten und der Leistung einer Person Rückschlüsse zu ziehen. In der Vergangenheit wurden viele Theorien und Modelle aufgestellt, die uns die Struktur der Intelligenz abbilden sollen. Bis heute ist es den Wissenschaftlern wegen der Undeterminiertheit nicht gelungen, eine einheitliche Definition dieses Begriffes zu formulieren (Holling et al, 2004).

Schlägt man im Wörterbuch unter dem Begriff „Intelligenz“ nach, so findet sich bei Fröhlich (2002) folgende Erklärung. Der Ausdruck lässt sich aus den lateinischen Worten „intellego“ (erkennen, begreifen, verstehen) und „intellegentia“ (Einsicht, Verständnis, Kennerschaft) herleiten. Er bezeichnet „die generelle Fähigkeit des Menschen, durch die Erfassung von Bedeutungsbeziehungen, d.h. durch Einsicht und Denken, neue Aufgaben lösen oder neue Situationen bewältigen zu können“ (S.243).

Hofstätter (1957) beschreibt Intelligenz als „die den innerhalb einer bestimmten Kultur erfolgreichen gemeinsamen Fähigkeiten“ (Roth, 1998, S. 12).

Wechsler wiederum definiert Intelligenz als „die Fähigkeit, zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinanderzusetzen“ (Gittler & Arendasy, 2000, S.40).

Jedes Verfahren zur Messung der Intelligenz wurde auf der Grundlage einer bestimmten Rahmentheorie konstruiert. Um einen geeigneten Test für eine bestimmte Fragestellung

auswählen zu können, sowie die Ergebnisse korrekt interpretieren zu können, ist es wichtig diese Theorien, auf die ein Test beruht, genau zu kennen (Holling et al, 2004).

Nachdem es aufgrund der Vielzahl der bestehenden Modelle nicht möglich ist, alle zu beschreiben, soll hier ein kurzer Überblick über die wichtigsten Intelligenzstrukturmodelle dargestellt werden.

1.1. Die Zwei-Faktorentheorie bzw. die Generalfaktorentheorie von Spearman

Charles Spearman (1863 – 1945) entdeckte anhand verschiedener kognitiver Tests, dass die Leistungen von Versuchspersonen positiv miteinander korrelieren. Aufgrund dieser Beobachtungen kam er zu dem Schluss, dass diese kognitiven Leistungen auf einen gemeinsamen Faktor zurückzuführen seien, den Spearman als Generalfaktor (g) bezeichnet. So formulierte der britische Forscher 1904 erstmals eine explizite Modellvorstellung über das Konstrukt der Intelligenz unter Anwendung der von ihm entwickelten Korrelationsrechnung sowie der Faktorenanalyse. Nachdem die Zusammenhänge zwischen den Tests nie perfekt waren, ging er weiter von der Annahme aus, dass die bisher nicht erklärte Varianz auf zusätzliche spezifische Faktoren begründet ist. Spearman unterscheidet somit zwei Faktoren, welche die Leistungen in kognitiven Tests definieren sollen. Einerseits nimmt er eine allgemeine Form der Intelligenz an, die er durch den g-Faktor beschreibt und andererseits soll es einen spezifischen Faktor geben (s-Faktor), der die typischen Fähigkeiten der jeweiligen Aufgaben abbildet. Daher spricht man in diesem Zusammenhang auch von der Zwei-Faktoren-Theorie oder auch Generalfaktorentheorie.

Die Erklärung der Varianz anhand dieser zwei Faktoren schien allerdings nicht zufriedenstellend zu sein, da empirisch gezeigt werden konnte, dass manche Aufgabentypen einen stärkeren Zusammenhang aufweisen, als durch ihre jeweiligen Ladungen auf g zu

erwarten wäre. Auch die Analysen der spezifischen Faktoren wiesen Schwachpunkte auf. Zwischen den einzelnen s-Faktoren konnten Restkorrelationen gefunden werden, die Spearman weiter als „spezielle Generalfaktoren“ charakterisierte.

Tests, die anhand dieser Modellvorstellung entwickelt wurden, sind beispielsweise die von Wechsler konstruierten Verfahren HAWIK-III und HAWIK-R. (Holling et al, 2004)

1.2. Die Primärfaktorentheorie von Thurstone

Louis Leon Thurstone (1887 – 1955) geht, entgegen der Meinung von Spearman, von der Annahme aus, dass bei der Lösung kognitiver Aufgaben mehrere nebeneinander stehende Gruppenfaktoren (primary mental abilities) in jeweils unterschiedlicher Gewichtung beteiligt sind. Allerdings wird keiner dieser Faktoren als allgemein betrachtet, sodass er bei jeder kognitiven Leistung beteiligt ist. Daher ist es nach Thurstones Auffassung nicht möglich einen Wert für die Intelligenz, wie eben den g-Faktor nach Spearman, zu berechnen. Der Psychologe Thurstone war maßgeblich an der Entwicklung der multiplen Faktorenanalyse beteiligt und fand zunächst in seinen Analysen neun, später dann sieben grundlegende Faktoren der Intelligenz, anhand derer man ein Profil der Ausprägungsgrade erstellen kann (Holling et al, 2004; Süllwold, 1977; Jäger, 1973).

Folgende primäre Fähigkeiten lassen sich demnach identifizieren:

1. *verbal comprehension* („Kenntnis von Wörtern und ihrer Bedeutung sowie deren angemessene Verwendung im Gespräch.“)
2. *word fluency* („rasches Produzieren von Wörtern, die bestimmten strukturellen oder symbolischen Erfordernissen entsprechen.“)
3. *number* („Geschwindigkeit und Präzision bei einfachen arithmetischen Aufgaben.“)

4. *space* („Bewältigung von Aufgaben, die räumliches Vorstellen und Orientieren sowie das Erkennen von Objekten unter anderem Bezugswinkel erfordern.“)
5. *memory* („Behalten paarweise gelernter Assoziationen.“)
6. *perceptual speed* („Geschwindigkeit beim Vergleich oder der Identifikation visueller Konfigurationen.“)
7. *induction oder reasoning, general* („Schlussfolgerndes Denken im Sinne des Auffindens einer allgemeinen Regel in einer vorgegebenen Abfolge von Zahlen oder Symbolen und Anwendung derselben bei der Vorhersage des nächstfolgenden Elementes.“) (Thurstone & Thurstone, 1941, zitiert nach Amelang et al, 2006, S. 181)

Diese Fähigkeiten sind nach Thurstones Meinung zwar unterschiedlich, doch in gewisser Weise voneinander abhängig (Holling et al, 2004).

Thurstone entdeckte noch zwei weitere, nicht näher interpretierbare Faktoren. Diese waren einerseits die Fähigkeit logische Probleme zu lösen (*reasoning*) und andererseits ein Faktor, der anhand deduktiver Eigenschaften beschrieben werden kann (Roth, 1998).

Der Grund für die unterschiedlichen Ergebnisse der Studien von Spearman und Thurstone ist vor allem in der methodischen Vorgehensweise zu finden. In Thurstones Studien wurden vorwiegend sehr heterogene Tests verwendet, wodurch auch kleinere Korrelationen zwischen den Testleistungen beobachtet wurden. Weiters waren hauptsächlich Studenten in seiner Stichprobe vertreten, weshalb deutlich unter dem Durchschnitt liegende Leistungen kaum vertreten waren. Die geringe Leistungsvarianz führt zusätzlich zu niedrigeren Korrelationen. Verfahren zur Messung der Intelligenz, die auf der Grundlage von Thurstones Primärfaktoren beruhen sind beispielsweise alle Versionen der Intelligenz-Struktur-Tests von Amthauer (I-S-T; I-S-T 70; I-S-T 2000 R). Entgegen der Auffassung von Thurstone wird hier letztendlich doch die Möglichkeit zur Berechnung einer allgemeinen Intelligenz vorgeschlagen (Holling et al, 2004).

1.3. Das morphologische Modell von Guilford

Aufgrund der immer weiter vorangetriebenen Differenzierung der Intelligenzfaktoren resultierte eine große Menge unterschiedlichster Intelligenzkomponenten. Guilford war daher bestrebt diese riesige Anzahl in eine gewisse Ordnung zu bringen und versuchte die kognitiven Leistungen anhand dreier unabhängiger Dimensionen zu erklären. So entstand wohl eines der umfassendsten Modelle der Strukturforschung (Guilford & Hoepfner, 1976; Jäger, 1973).

1) Operationen

- Kognition: das Finden und Erkennen von Informationen,
- Gedächtnis: das Behalten von und Verfügen über Information,
- Divergente Produktion: das Bearbeiten von Problemen mit mehrdeutigen Lösungen auf nicht konventionelle Weise,
- Konvergente Produktion: das Bearbeiten von Aufgaben mit eindeutigen, anerkannten Lösungen,
- Bewertung: die Beurteilung des Erkannten oder Produzierten hinsichtlich seiner Richtigkeit und Brauchbarkeit;

2) Inhalte

- Figurale: alle - nicht nur optisch – wahrnehmbaren Gestalten,
- Symbolische: Zahlen, Buchstaben, Worte (ohne Berücksichtigung ihrer Bedeutung),
- Semantische: (zunächst sprachliche) Bedeutungen,
- Verhaltensmäßige: angenommene, aber noch nicht näher erforschte, in Beziehung zu „sozialer Intelligenz“ stehende Inhalte;

3) Produkte

- Einheiten, Klassen, Relationen, Systeme, Transformationen und Implikationen.

(Roth, 1998, S. 25)

Aufgrund der fünf Operationen, der vier Inhaltsdimensionen und der sechs Produktkategorien ergeben sich somit 120 Kombinationsmöglichkeiten, die jeweils einen unabhängigen Intelligenzfaktor darstellen (Guilford & Hoepfner, 1976).

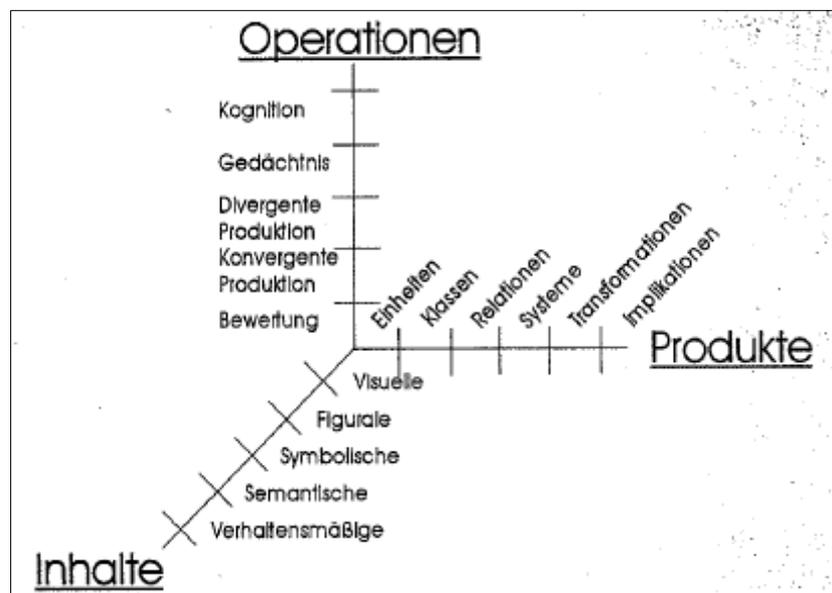


Abbildung 1: „Strukturmodell der Intelligenz nach Guilford“ (Roth, 1998, S. 25)

1.4. Die Theorie der fluiden und kristallinen Intelligenz von Cattell

Raymond B. Cattell, ein Schüler Spearman, stellte 1957 ein Modell der Intelligenz vor, das durch zwei generelle Faktoren definiert werden kann. Cattell geht in dieser Theorie von zwei Ebenen aus. Auf der unteren Ebene befinden sich, ähnlich dem Modell von Thurstone, unterschiedliche, aber miteinander in Zusammenhang stehende primäre Faktoren. Auf einer hierarchisch höher gelegenen Ebene sind zwei generelle Faktoren angesiedelt, die als fluide (gf) und kristalline (gc) Intelligenzfaktoren bezeichnet werden.

Ersterer beschreibt die von Geburt an angelegte Fähigkeit einer Person, sich neuartigen Problemen zu stellen, sowie sich in neuen Situationen zurecht zu finden. Diese Dimension ist unabhängig vom Wissensstand sowie von kulturellen und gesellschaftlichen Einflüssen.

(Holling et al, 2004) Sie wird vor allem durch „Induktives Schließen“, „Figurale Beziehungen“ und „Intellektuelle Geschwindigkeit“ gemessen.

Die kristallinen Fertigkeiten hingegen basieren auf Lernerfahrungen und Wissensbeständen, die ständig weiterentwickelt werden. Zur Messung wird hier überwiegend das „Verbale Verständnis“ oder „Mechanische Kenntnisse“ herangezogen (Horn, 1986, zitiert nach Holling et al, 2004, S. 21).

Für die kristalline und fluide Intelligenz kann ein unterschiedlicher Entwicklungsprozess festgestellt werden. Nach einem steilen Anstieg um das 25. Lebensjahr bleibt die kristalline Intelligenz bis ins hohe Alter recht stabil und kann oft noch einen weiteren Anstieg verzeichnen, während die fluide Intelligenz mit zunehmendem Alter allmählich absinkt (Süß, 2003).

Cattell geht weiter von der Annahme aus, dass fluide Intelligenz notwendig ist, um kristalline Intelligenz erwerben zu können. Er entwickelte ein sprachfreies Verfahren, um die fluide Intelligenz möglichst kulturfair erfassen zu können.

Aufgrund der Erkenntnis, dass beide Faktoren (gf und gc) miteinander korrelieren erweitert Cattell seine Theorie um den Faktor gf(h), der als Faktor dritter Ordnung eingeführt wird und einen größeren Zusammenhang zu gf aufweist. Das Modell kann allerdings faktorenanalytisch nicht überprüft werden, da das Auffinden von Faktoren dritter Ordnung mathematisch nicht zu lösen ist (Jensen, 1998, zitiert nach Holling et al, 2004).

1.5. Das Berliner Intelligenzstrukturmodell (BIS) von Jäger

Das von Jäger 1982 entwickelte Intelligenzstrukturmodell stellt eine Vereinigung der Erkenntnisse der Theorien von Spearman, Thurstone und Guilford dar (Amelang, 2006; Süß, 2003). Es kann als integratives Modell aufgefasst werden, welches eine breite Palette an intellektuellen Fähigkeiten abdeckt. Als Basis seiner Theorie dient ein sehr umfangreiches Inventar an Kreativitäts- und Intelligenztests, welche seit den 70er Jahren in der testpsychologischen Diagnostik angewendet wurden. Nach Jägers Auffassung ist die höchste Ebene, nämlich die allgemeine Intelligenz, aus sieben verschiedenen Fähigkeiten zusammengesetzt, die weiter in operative und inhaltliche Fähigkeiten unterteilt werden können. Auf der dritten Ebene befinden sich somit 12 Zellen, die sich durch die Kombinationen der sieben Fähigkeiten ergeben (Holling et al, 2004; Süß, 2003).

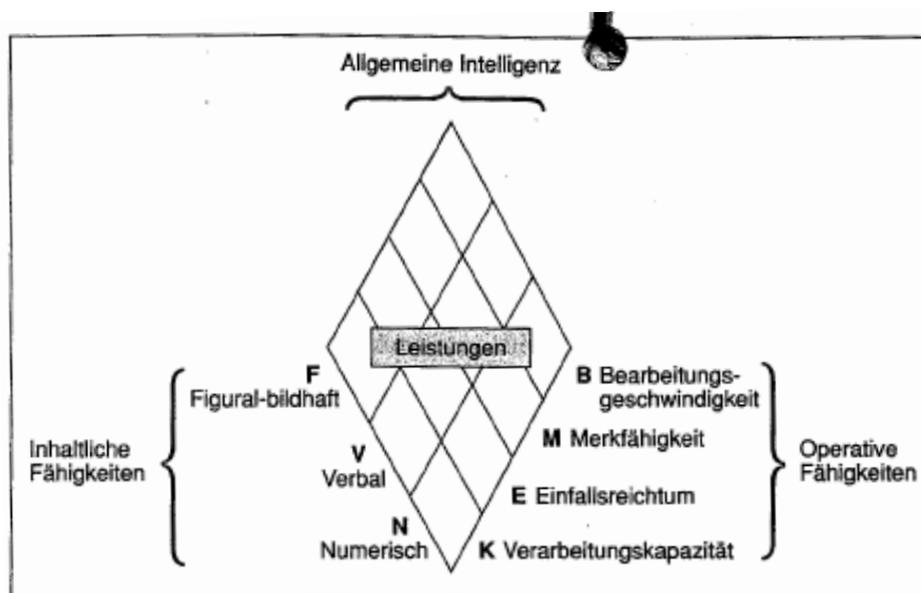


Abbildung 2: „Berliner Intelligenzstrukturmodell nach Jäger 1982, 1984“ (Holling et al, 2004, S. 24)

Operative Fähigkeiten:

- B *Bearbeitungsgeschwindigkeit*: Arbeitstempo, Auffassungsleichtigkeit und Konzentrationskraft beim Lösen einfach strukturierter Aufgaben von geringem Schwierigkeitsniveau.
- M *Merkfähigkeit*: Aktives Einprägen und kurzfristiges Wiedererkennen oder Reproduzieren von Informationen
- E *Einfallreichtum*: Flexible Produktion von vielfältigen Ideen und Lösungen für eine vorgegebene Problemstellung; erforderlich ist hierfür die Verfügbarkeit vielfältiger Informationen, ein Reichtum an Vorstellungen und das Sehen vieler verschiedener Seiten, Varianten, Gründe und Möglichkeiten von Gegenständen und Problemen.
- K *Verarbeitungskapazität*: Verarbeitung komplexer Informationen bei Aufgaben, die nicht auf Anhieb zu lösen sind, sondern Heranziehen, vielfältiges Beziehungsstiften, formallogisch exaktes Denken und sachgerechtes Beurteilen von Informationen erfordern.

Inhaltsgebundene Fähigkeiten:

- F *Anschaunungsgebundenes, figural-bildhaftes Denken*: Fähigkeit zum Umgang mit Aufgabenmaterial, dessen Bearbeitung figural-bildhaftes und/oder räumliches Vorstellen erfordert.
- V *Sprachgebundenes Denken*: Grad der Aneignung und der Verfügbarkeit des Symbolsystems Sprache.
- N *Zahlengebundenes Denken*: Grad der Aneignung und der Verfügbarkeit des Symbolsystems Zahlen. (Holling et al, 2004, S.24)

Zwei Verfahren, die auf den Erkenntnissen von Jäger aufbauen, sind der BIS-4 von Jäger, Süß und Beauducel (1997) sowie der BIS-HB von Jäger et al (2004) (Holling et al, 2004).

1.6. Die Three-Stratum-Theory von Carroll

Carroll stellte 1993 eine Theorie der Intelligenz auf, dessen Grundlage die Reanalyse sämtlicher in der Literatur beschriebener Korrelationsmatrizen war. (Süß, 2003)

Mit der Methode der Faktorenanalyse formulierte er ein hierarchisches Konzept mit 3 Ebenen, die er als narrow (Stratum I), broad (Stratum II) und general (Stratum III) bezeichnet (Carroll, 1993).

In der Hierarchie ganz oben angeordnet befindet sich das Stratum III, welches eine allgemeine Form der Intelligenz darstellt. Es ist durch eine hohe Generalität gekennzeichnet, die alle kognitiven Leistungen beeinflusst.

Auf der zweiten Ebene befinden sich acht Fähigkeiten mittelgradiger Spezialisierung:

1. *Fluide Intelligenz*: Basale Prozesse des schlussfolgernden, logischen Denkens und andere mentale Aktivitäten, die nur minimal durch Lernen und Akkulturation beeinflusst werden.
2. *Kristalline Intelligenz*: Mentale Prozesse, die nicht nur von der fluiden Intelligenz, sondern auch von Erfahrung, Lernen und Akkulturation abhängen.
3. *Allgemeine Gedächtnisfähigkeit*: Fähigkeit zum Lernen und Behalten neuer Inhalte oder neuen Verhaltens.
4. *Visuelle Wahrnehmung*: Fähigkeit, die bei allen Aufgaben eine Rolle spielt, die die Wahrnehmung visueller Formen erfordern.
5. *Auditive Wahrnehmung*: Fähigkeit zur Wahrnehmung oder Unterscheidung auditiver Klangmuster oder gesprochener Sprache.
6. *Abruffähigkeit*: Fähigkeit, Konzepte, Inhalte oder Ideen aus dem Langzeitgedächtnis abzurufen.

7. *Kognitive (Verarbeitungs-)Geschwindigkeit*: Geschwindigkeit der kognitiven Verarbeitung von Informationen.
8. *Entscheidungsgeschwindigkeit*: Entscheidungsgeschwindigkeit in verschiedenen Reaktionsaufgaben (Holling et al, 2004, S. 28-29).

Auf der Stratum I-Ebene sind wiederum 68 sehr spezifische Konstrukte angesiedelt, die von mehreren Stratum II-Fähigkeiten beeinflusst werden. Diese 68 Fähigkeiten stellen ziemlich alle Fähigkeitskonstrukte der Forschungsgeschichte dar (Süß, 2003).

Im Bereich der verbalen Fähigkeiten können laut Carroll (1993) folgende Unterscheidungen getroffen werden:

- Language Development
- Verbal / Printed Language
- Lexical Knowledge
- Reading Comprehension
- Reading Decoding
- Reading Speed
- Cloze Ability
- Spelling Ability
- Oral Language Proficiency
- Phonetic Coding
- Foreign Language Aptitude
- Communication Ability
- Listening Ability
- Oral Production
- Oral Style
- Writing Ability

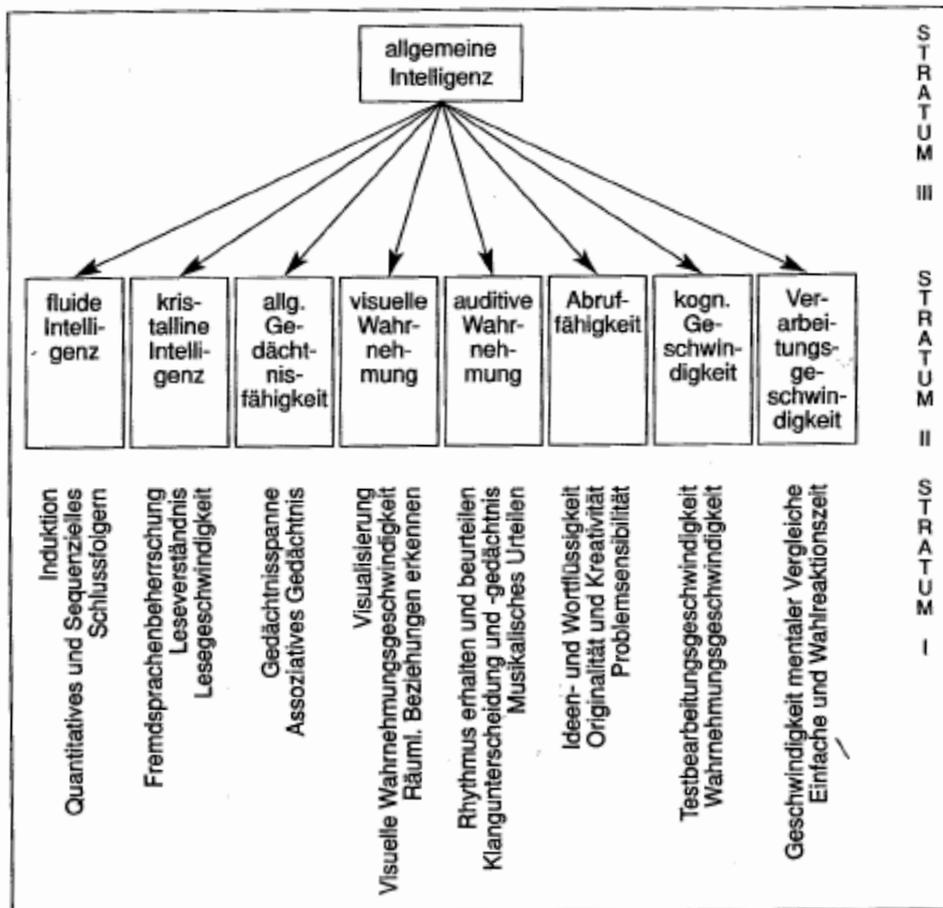


Abbildung 3: „Three-Stratum-Theorie nach Carroll 1993“ (Holling et al, 2004, S. 28)

1.7. Die Theorie der multiplen Intelligenzen von Gardner

Howard Gardner definiert Intelligenz als „biopsychologisches Potential zur Verarbeitung von Informationen, das in einem kulturellen Umfeld aktiviert werden kann, um Probleme zu lösen oder geistige oder materielle Güter zu schaffen, die in einer Kultur hohe Wertschätzung genießen“ (Gardner, 2002, S. 46). Er kritisierte die Annahme vieler Wissenschaftler, dass die Vollendung der geistigen Entwicklung eines Menschen hauptsächlich an deren akademischen Laufbahn, wie beispielsweise die eines Molekularbiologen, gemessen wurde. Aufgrund seines musikalischen und künstlerischen Interesses war er davon überzeugt, dass aus entwicklungspsychologischer Sicht den Fähigkeiten von Malern, Schriftstellern, Musikern, Tänzern und allen anderen Künstlern mehr Aufmerksamkeit entgegen gebracht werden sollte

(Gardner, 2002). Dies ist ein wesentlicher Aspekt, den Gardner in seiner Theorie der multiplen Intelligenzen berücksichtigt hat. Er versuchte, eine möglichst breite Palette an relativ unabhängigen Fähigkeiten in sein Modell einzubeziehen, die er anhand folgender Kriterien identifizieren konnte:

- Isolierung durch Hirnverletzungen
- Wunderkinder (Kinder, die in manchen Bereichen überdurchschnittliche Leistungen erzielen)
- Identifizierung von Kernoperationen (Annahme grundlegender Datenverarbeitungsmechanismen für spezifische Inputs)
- Abgrenzbare ontogenetische Entwicklungen
- Evolutionsgeschichtliche Hinweise
- Experimentalpsychologische Befunde
- Ergebnisse der Psychometrie
- Tendenz zur Verschlüsselung symbolischer Systeme (z.B. Sprache, bildliche Darstellung, Mathematik). (Gardner, 1991)

Nach Gardner können aufgrund seiner Kriterien sieben Formen der Intelligenz beschrieben werden:

- 1) *Sprachliche Intelligenz*: Sensibilität für geschriebene und gesprochene Sprache; Fähigkeit die Sprache zweckmäßig einzusetzen.
- 2) *Logisch-mathematische Intelligenz*: Logische Analyse von Problemen, Durchführung mathematischer Operationen und Untersuchung wissenschaftlicher Fragestellungen.
- 3) *Musikalische Intelligenz*: Begabung zum Musizieren und Komponieren, sowie ein Gefühl für Prinzipien der Musik
- 4) *Körperlich-kinästhetische Intelligenz*: Fähigkeit den gesamten Körper sowie einzelne Körperteile zur Lösung eines Problems oder zur Kreation von Produkten einzusetzen.

- 5) *Räumliche Intelligenz*: Erfassen der Struktur weiter Umgebungen, wie z.B. beim Kartenlesen aber auch das Erfassen enger Raumgrenzen, wie es etwa für Chirurgen, Bildhauer oder Architekten von Bedeutung ist.
- 6) *Interpersonale Intelligenz*: Das Erkennen von Absichten, Motiven und Wünschen anderer Personen sowie der Herstellung einer entsprechenden Kooperation.
- 7) *Intrapersonale Intelligenz*: Die Begabung einer Person die eigenen Wünsche, Ängste und Fähigkeiten, sprich die eigene Persönlichkeit, realistisch wahrzunehmen und im Alltag gezielt einzusetzen.

Howard Gardner weist darauf hin, dass sein Modell keineswegs eine absolute Darstellung der Intelligenzstruktur bietet. Allerdings werden hier Dimensionen berücksichtigt, die bis dato in der Intelligenzforschung kaum von Bedeutung waren (Gardner, 2002).

2. Intelligenzdiagnostik und Diagnostik verbaler Fähigkeiten

2.1. Die Anfänge der Intelligenzdiagnostik



Wenn man über die Geschichte der Intelligenzforschung liest, wird man auf jeden Fall mit den Namen Alfred Binet und Theodore Simon, die weltweit den ersten brauchbaren Intelligenztest entwickelten, konfrontiert. Alfred Binet wurde 1857 in Nizza als Sohn einer Künstlerin und eines Arztes geboren und beendete seine erste akademische Ausbildung 1878 als Jurist. Später wandte er sich den Naturwissenschaften und der Medizin zu. Sein großes Interesse galt der Psychologie, die damals noch keine eigenständige akademische Disziplin darstellte. Er gründete gemeinsam mit Henri Beaunis 1889 das erste psycho-physiologische Laboratorium an der Sorbonne. Im Jahr 1899 kam Theodore Simon im Alter von 26 Jahren, zu Binet und bat ihn, als Mitarbeiter aufzunehmen. Aufgrund seiner Praktikumsstelle in einem Heim für zurückgebliebene Erwachsene und Kinder hatte er leichten Zugang zu möglichen Testpersonen und gewann somit Binets Zustimmung.

Aufgrund der einheitlichen Beschulung aller Kinder, unabhängig von deren Leistungsniveau, kam es zu enormen Schwierigkeiten innerhalb des Schulunterrichtes. Diese Tatsache bewegte Binet und Simon dazu, einen Test zu entwickeln, der das Leistungsniveau der Kinder identifizieren sollte, um eine angepasste Beschulung zu ermöglichen (Pflaum, 2000; Funke, 2006).

Binet widersetzte sich der Annahme, dass Intelligenz aus einzelnen elementaren Fähigkeiten wie Sinnesschärfe, Merkfähigkeit, Reaktionszeit und ähnliches zusammengesetzt ist. Er

verstand Intelligenz als „eine a priori komplexe und daher entsprechend zu messende Fähigkeit“ (Strube, 1977, S. 2).

Wie Binet sich die Entwicklung eines solchen Tests im Detail vorstellt, soll nun anhand des folgenden Zitates verdeutlicht werden:

Man denkt eine große Menge von Proben aus, die schnell ausgeführt werden können und zugleich genau sind; ferner sollen sie eine ansteigende Schwierigkeit darstellen; dann macht man mit diesen Proben an einer großen Zahl von Kindern verschiedenen Alters Untersuchungen, um sie zu probieren; man notiert die Resultate; man sucht dann diejenigen Proben heraus, die sich für ein gegebenes Alter ergeben und die jüngere Kinder – wobei es sich nur um ein Jahr handeln würde – im Durchschnitt nicht bestehen können; dann stellt man auf die Art eine Stufenleiter für die Intelligenz auf, mit Hilfe derer man bestimmen kann, ob eine Person, um die es sich handelt, die Intelligenz ihres Alters hat, oder ob sie im Rückstand oder im Vorsprung ist, und wie viele Monate oder Jahre dieser Vorsprung oder Rückstand beträgt (Anschütz & Ruttmann 1912, S. 105).

Binets Töchter Madeleine und Alice waren die ersten Personen, die zur Testung herangezogen wurden. Die erste Version des Tests erschien 1905 und beinhaltete 30 Aufgaben, die anhand von Studien an Elementarschulen und Mutterschulen, aber auch in Anstalten für „Idioten“ erprobt wurden (Funke, 2006; Anschütz & Ruttmann, 1912).

Überblick der Aufgabentypen:

1. Das Sehen (Koordination zwischen Augenbewegung und Kopf)
2. Durch taktile Bewegung ausgelöstes Greifen
3. Durch visuelle Wahrnehmung ausgelöstes Greifen
4. Bekanntheit von Nahrung

5. Suche nach Nahrung, erschwert durch ein kleines physisches Hindernis
6. Ausführung einfacher Anweisungen und Nachahmung einfacher Handbewegungen
7. Verbales Beschreiben von Objekten
8. Verbales Beschreiben von Bildern
9. Benennung ausgewiesener Objekte
10. Sofortiger Vergleich zweier Linien unterschiedlicher Länge
11. Wiederholung dreier Ziffern
12. Vergleich zweier Gewichte
13. Beeinflussbarkeit
14. Verbale Definition bekannter Objekte
15. Wiederholung von aus 15 Wörtern zusammengesetzten Sätzen
16. Unterschiede zwischen mehreren bekannten Objekten in der Erinnerung
17. Anwendung des Gedächtnisses bei Bildern
18. Zeichnen aus dem Gedächtnis
19. Unverzögliche Wiederholung von Ziffern
20. Ähnlichkeiten zwischen mehreren bekannten Objekten in der Erinnerung
21. Vergleich von Längen
22. Anordnen von fünf Gewichten
23. Lücken von Gewichten
24. Finden von Reimen
25. Auffüllen verbaler Lücken
26. Einordnung dreier Worte in einen sie beinhaltenden Satz
27. Antwort auf eine abstrakte Frage
28. Umkehrung des Uhrzeigersinnes
29. Zerschneiden

30. Definition abstrakter Begriffe (Funke, 2006, S. 29 – 32)¹

Für all diese Aufgabentypen wurde eine genaue Angabe gemacht, wie diese durchzuführen seien, um eine möglichst standardisierte Darbietung zu gewährleisten. Auch mögliche falsche Antworten wurden genauestens analysiert und eingestuft. Aufgrund dieses Tests sollte es nun möglich sein, das Intelligenzalter eines Kindes zu ermitteln. Nachdem für jede Altersgruppe fünf einzelne Aufgaben erstellt wurden, wird für jede richtige Lösung 1/5 Jahr berechnet und für jede halbrichtige Antwort 1/10 Jahr. Ausgangspunkt für die Berechnung ist jene Altersgruppe, in der zu letzt alle fünf Aufgaben richtig beantwortet wurden. Alle weiteren richtigen Aufgaben werden anhand der oben dargestellten Verrechnungsweise addiert. Hat beispielsweise ein Kind alle Aufgaben in der Altersgruppe der fünfjährigen letztmals richtig gelöst, so geht man zunächst von einem Intelligenzalter von fünf Jahren aus. Hat das Kind in den folgenden Jahresreihen vier Aufgaben vollständig und sechs weitere teilweise richtig gelöst, so ergibt sich folgende Addition: $5 + 4/5 + 6/10 = 6,4$. In diesem Beispiel würde das Kind ein Intelligenzalter von 6,4 Jahren erreichen (Probst, 1960).

Die spätere Entwicklung des Intelligenzquotienten (IQ) beruht auf diesem Ansatz, indem man das Verhältnis von Intelligenz-Alter zum Lebensalter betrachtet. Aufgrund der nun möglichen Leistungsdifferenzierung wurde 1909 eine gesetzliche Grundlage für die Aufnahme spezieller Schulklassen für Kinder mit einer Lernbehinderung geschaffen. Der Simon-Binet-Test wurde nicht nur in Frankreich erfolgreich eingesetzt, sondern verbreitete sich rasch in Deutschland und den USA.

Seit der ersten Veröffentlichung des Intelligenztests von Binet und Simon sind nun schon mehr als einhundert Jahre vergangen und man sollte glauben, dass kaum noch etwas von den damaligen Aufgabentypen und Methoden Gültigkeit besitzt. Diese Annahme ist vielleicht für

¹ Eine genauere Beschreibung der Aufgaben kann ebenda gefunden werden.

manche Testaufgaben zutreffend, doch was die methodische Vorgehensweise betrifft, hat sich kaum etwas verändert:

1. Steigerung der Reliabilität durch die Vorgabe mehrerer Items
2. Staffelung nach dem Schwierigkeitsgrad
3. Standardisierung durch genaue Instruktion
4. Normierung
5. Konzentration auf ein konkretes Leistungsverhalten
6. als Grundlage der Beurteilung dient nicht alleine das Testergebnis (Funke, 2006)

2.2. Dimensionen sprachlicher Fähigkeiten

Wie im Kapitel über die Intelligenzstrukturmodelle zu sehen ist, gibt es unterschiedliche Auffassungen darüber, wie das Konstrukt der Intelligenz zusammengesetzt ist. Viele Theorien gehen von der Annahme aus, dass die verbalen Fähigkeiten einen wesentlichen Aspekt intellektuellen Verhaltens darstellen. Die meisten modernen Intelligenztestbatterien beinhalten Aufgaben, um die sprachliche Intelligenz eines Menschen zu analysieren. Weiters gibt es einige spezifische Tests, die sich ausschließlich mit der Erfassung verbaler Fähigkeiten befassen. Wie die Operationalisierungsversuche sprachlicher Fähigkeiten zusammengefasst werden können, soll nun das folgende Kapitel erläutern.

Aufgrund der Analyse eines sehr umfangreichen Testmaterials beschreibt Acker (2001) vierzehn Dimensionen verbal-intelligenter Fähigkeiten, die hier kurz vorgestellt werden sollen. Alle Angaben zum folgenden Kapitel (2.2.) sind dieser Literatur entnommen, ausgenommen Verweise auf andere Quellen.

| | | | | |
|---|--|--|---|--|
| ANA DTD (AF) I-S-T 70 (AN) KFT 4-13 ⁺ (V4) M-I-T (WV) SASKA (An) W-I-T (AL) | ANT SASKA (At) | ASS VKT (VS, NE, GE, UV, SE) | KLA HAWIE-R (GF) I-S-T 70 (GE) KFT 4-13 (V3) SASKA (KI) WIP 86 (GF) | KSI M-I-T (SW) W-I-T (SW) I-S-T 70 (SE) KFT 4-13 ⁺ (V2) DTD (IV) |
| MEK DTD (LV) | REK DTD (PW) DTD (TS) LPS (s5) PSB (s1+2) | SEL I-S-T 70 (WA) SASKA (Se) | UNM M-I-T (UN) | VOR VKT (US) |
| WEK MWT-A/B WST | WV VKI | FLU DTD (AW) LPS (s6) PSB (s5) W-I-T (WG) | FLU ^{ASS} VKT (WA, WE) | SYN AID (SF) CFT-20 (ET) HAWIE-R (WT) KFT 4-13 (V1) M-I-T (WB) SASKA (Sy 1-3) VKT (ÄK) W-I-T (GW) |

Abbildung 4: „Dimensionen verbal-intelligenter Fähigkeiten nach Acker 2001“

2.2.1. Analogien (ANA)

Diese Dimension befasst sich mit dem Finden von Wörtern mit gleicher oder ähnlicher Bedeutung. Dieser Aufgabentyp ist in der Testliteratur weit verbreitet und ist beispielsweise in folgenden Verfahren vertreten: DTD (AF), I-S-T 70 (AN), KFT 4-13⁺ (V4), M-I-T (WV), SASKA (An), W-I-T (AL)

Beispiel aus dem KFT 4-13⁺ (V4 - Wortanalogien):

Im Kognitiven Fähigkeitstest von Heller, Gaedike und Weinläder (1985) wird ein Wortpaar vorgegeben „dessen Testteile in einem bestimmten Verhältnis zueinander stehen. Zu einem dritten Wort ist aus fünf Antwortalternativen diejenige herauszufinden, die mit dem dritten Wort in gleicher Relation (Analogie) steht wie die beiden ersten.“ (Heller, Gaedike, Weinläder, 1985, S. 7, zitiert nach Acker, 2001, S. 36)

Säure : Lauge → These :

A Axiom B Antithese C Synthese D Aspekt E Argument

2.2.2. Antonyme (ANT)

Diese ist eine relativ gering verbreitete Aufgabenart, bei der zu einem Ausgangsbegriff ein Gegenteiliges Wort aus einer vorgegebenen Reihe gefunden werden muss. Hierzu gibt Acker (2001) an, lediglich einen Test gefunden zu haben.

Beispiel aus dem SASKA Subtest 2 (At)

Die Instruktion zum Subtest 2 im SASKA von Riegler (1967) lautet:

Es muss „auf der gleichen, durch das Abstraktionswort angegebenen Abstraktionsebene verblieben werden, nur daß in diesem Fall die Begriffe im entgegengesetzten Sinn als ‚gleichbedeutend‘ zu erkennen sind, d.h. zu dem links stehenden Testwort ist ein zweites Wort unter den fünf Auswahlantworten zu finden, welches das Gegenteil des Testwortes bedeutet.“ (Riegel, 1967, S.7, zitiert nach Acker, 2001, S. 40)

Ermüdung ... a) Aufwachen b) Erholung c) Genesung d) Einschlafen e) Ansporn

2.2.3. Assoziationen (ASS)

Diese Dimension befasst sich mit den assoziativen Aspekten verbaler Fähigkeiten, wobei diese sich nur schwer isoliert erfassen lassen. Assoziationen und verbal produktive Fähigkeiten werden oft vermischt und die quantitative Leistung tritt hervor, wodurch die ursprüngliche Intention verloren geht. Vielleicht ist auch das genau der Grund warum Acker (2001) nur ein Testinstrument gefunden hat, dass diese Fähigkeiten zu erfassen versucht.

Beispiel aus dem VKT Subtest 3 (VS – Vierwort-Sätze):

Die Instruktion zum Subtest 3 des Verbalen Kreativitätstests von Schoppe (1975) lautet:

Es werden Ihnen vier Anfangsbuchstaben vorgegeben, die Sie nach Belieben als groß- oder kleingeschrieben ansehen können. Zu jedem dieser vier Anfangsbuchstaben soll Ihnen je ein Wort einfallen, daß diese vier Wörter dann einen Vierwort-Satz ergeben. Es dürfen keine Füllwörter eingeschoben werden. Die Reihenfolge der Anfangsbuchstaben darf aber verändert werden [...]. Von Ihnen wird in erster Linie erwartet, daß Sie möglichst viele verschiedene Vierwort-Sätze bilden können.
(Schoppe, 1975, S. 3, zitiert nach Acker, 2001, S. 42)

R-S-A-M

Rosie surft am Meer

Rudi sitzt am Misthaufen. ...

2.2.4. Klassifikation (KLA)

Diese Dimension beinhaltet Aufgaben, die Leistungen bezüglich der Klassifikation von sprachlichem Material erfordern, wobei hier zwischen aktiver und passiver Klassifikationsleistung unterschieden werden kann. Tests, die diesen Aufgabentyp beinhalten sind beispielsweise der HAWIE-R (GF), I-S-T 70 (GE), KFT 4-13 (V3), SASKA (KI) und der WIP 86 (GF)

Beispiel aus dem HAWIE-R Subtest 11 (Gemeinsamkeiten finden - GF) – aktive

Klassifikation

Beim Hamburg Wechsler Intelligenztest für Erwachsene in der Revision von Tewes (1991) werden einer Testperson zwei Begriffe genannt, für die sie selbstständig, d.h. ohne jegliche Vorgaben von Auswahlmöglichkeiten, eine Gemeinsamkeit finden soll.

Was ist das Gemeinsame bei...?

Apfelsine – Banane

2.2.5. Komplexe sprachliche Inhalte (KSI)

Testitems zu dieser Dimension erfordern die Kenntnis und Anwendung komplexer, sprachlicher Inhalte. Sie versuchen das Verständnis ganzer Sätze auf verschiedene Art und Weise zu messen. Verfahren zur Erfassung dieser Fähigkeiten sind: M-I-T (SW), W-I-T (SW), I-S-T 70 (SE), KFT 4-13⁺ (V2), DTD (VI)

Beispiel aus dem W-I-T Subtest 5 (Sprichwörter - SW):

Dieser Subtest des Wilde-Intelligenz-Test von Jäger und Althoff (1983) soll erfassen, ob eine Testperson die Bedeutung eines Satzes versteht und zwei weitere sinngemäße Sätze aus fünf Antwortalternativen finden kann.

Je mehr einer hat, um so mehr einer will

- a) Wem das Glück den Finger reicht, der soll ihm die Hand bieten
- b) Stiehlt man wenig, kommt man ins Gefängnis, stiehlt man viel, fährt man im Wagen
- c) Gib mir Raum zum Sitzen und ich werde mir welchen zum Niederlegen machen
- d) Wer gerne trägt, dem ladet jeder auf
- e) Wem mehr Freiheit gestattet ist, als vernünftig ist, der wird noch mehr begehren, als ihm gestattet ist

2.2.6. Merkfähigkeit (MEK)

Auch wenn diese Dimension in mehreren Subtests beteiligt ist, behandelt Acker (2001) diese als eigenständige Dimension, da die Merkfähigkeit beim nachfolgenden Test eine große Rolle spielt.

Beispiel aus dem DTD Subtest 5 (Leseverständnis - LV):

Im fünften Subtest des Diagnostischen Tests Deutsch von Nauck und Otte (1980) werden zwei kurze Geschichten präsentiert, wobei jeweils im Anschluss daran einige Fragen gestellt werden. Aus vier Antwortmöglichkeiten soll dann die richtige Lösung gefunden werden.

„Laut Autoren liegt der Schwerpunkt bei der ‚analytischen Sinnentnahme‘ sowie dem ‚inhaltlichen Erschließen zusammenhängender Texte‘. (Acker, 2001, S. 62)

Peter spielt mit dem Dackel seines größeren Bruders. Der Hund holt immer wieder die fortgeworfenen Stöckchen zurück. Plötzlich wittert der Dackel eine Katze. Mit wütendem Bellen stürmt er davon. Nach wenigen Minuten kommt er mit eingezogenem Schwanz zurück.

A

Was erfährst du über den Dackel in der Geschichte?

- a) Er ist krummbeinig
- b) Er kann Katzen töten
- c) Er jagt Katzen
- d) Er ist braun

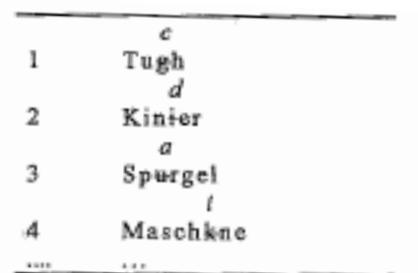
2.2.7. Rechtschreibkenntnisse (REK)

Obwohl manche Tests vorgeben andere Dimensionen zu prüfen, messen einige die Dimension der Rechtschreibkenntnisse. Tests, die hier angeführt werden können sind beispielsweise der DTD (Passiver Wortschatz), DTD (Textstrukturierung), LPS (s5) und das PSB (s1+2).

Beispiel aus dem DTD Subtest 1 (passiver Wortschatz - PW):

Der erste Subtest des Diagnostischen Tests Deutsch von Nauck und Otte (1980) listet eine Reihe von Worten auf, wobei jeweils ein Buchstabe pro Wort dieses sinnlos macht. Aufgabe der Testperson ist es nun diesen Buchstaben zu identifizieren, durchzustreichen und den richtigen Buchstaben darüber zu schreiben.

„Erfasst wird die Fähigkeit der Schüler, ihnen bekannte, jedoch verfremdet dargestellte Wörter zu identifizieren. Hierbei müssen die Schüler ihren passiven Wortschatz systematisch zur Entdeckung der Verfremdung einsetzen.“ (Nauck & Otte, 1980, S. 6, zitiert nach Acker, 2001, S. 64)



| | |
|---|----------|
| | <i>c</i> |
| 1 | Tugh |
| | <i>d</i> |
| 2 | Kinier |
| | <i>a</i> |
| 3 | Spurgel |
| | <i>t</i> |
| 4 | Maschkne |

Abbildung 5: „DTD Subtest 1 (Passiver Wortschatz)“ (Acker, 2001, S.65)

2.2.8. Selektion (SEL)

Subtests dieser Dimensionen verlangen von der Testperson Selektionsfähigkeiten im Bereich der Sprache. Gemeint sind hier beispielsweise Aufgaben, die anhand eines Ausgangsbegriffes auffordern, ein gänzlich unterschiedliches Wort, oder eines, das dieses am besten beschreibt,

aus vorgegebenen Alternativen zu wählen. Subtests mit diesem Aufgabentyp sind beispielsweise im I-S-T 70 (WA) und SASKA (Se) enthalten.

Beispiel aus dem I-S-T 70 Subtest 2 (Wortauswahl - WA)

Bei diesem Untertest des Intelligenz Struktur Tests von Amthauer (1970) werden der Testperson fünf Wörter, die gewisse Ähnlichkeiten aufweisen, vorgegeben. Die Aufgabe besteht nun darin, dasjenige Wort zu finden, das nicht zu den anderen passt. Die Autoren weisen darauf hin, dass dieser Subtest das „Erfassen von Sprachlichen Bedeutungsinhalten“, „Sprachgefühl“, „induktives sprachliches Denken“, „Einfühlungsfähigkeit“ und „rezeptive Komponenten“ erfasst. (Acker, 2001, S. 67)

24. a) Rede b) Gespräch c) Abstimmung d) Ansprache e) Diskussion

2.2.9. Unmöglichkeiten (UNM)

Diese Dimension versucht Sprachverständnis auf eine komplett andere Art und Weise zu messen als andere Testverfahren. Hier geht es um die Bewertung von Aussagen bezüglich ihrer Möglichkeit bzw. Unmöglichkeit. Der einzige Test mit dieser Dimension ist der M-I-T (Unmöglichkeiten).

Beispiel aus dem M-I-T Subtest 10 (Unmöglichkeiten – UN)

Jede Aufgabe des Subtests 10 des Mannheimer Intelligenztests von Conrad et al (1971) besteht aus fünf Behauptungen, die nach ihrer Richtigkeit überprüft werden sollen. Von den fünf Antwortalternativen ist immer eine richtig und die anderen falsch oder eine falsch und die anderen richtig.

4. Es ist unmöglich, dass eine Sackgasse
- a länger als 1000 m ist
 - b vor einem Haus endet
 - c breiter als 10 m ist
 - d durchfahren werden kann
 - e mit Kopfsteinpflaster gepflastert ist

2.2.10. Vorstellungsvermögen (VOR)

Anhand des VKT Subtest 8 konnte eine weitere Dimension sprachlicher Fähigkeiten identifiziert werden. Hier handelt es sich um die Anwendung verbal-produktiver Fähigkeiten.

Beispiel aus dem VKT Subtest 8 (Utopische Situationen – US)

Der Subtest 8 des Verbalen Kreativitätstest von Schoppe (1975) erfordert von den getesteten Personen Leistungen im Bereich des Vorstellungsvermögens, wobei hier die quantitativen Aspekte wiederum nicht außer Acht gelassen werden können. Die Instruktion lautet:

Es wird Ihnen in einem kurzen Satz eine ungewöhnliche Situation vorgestellt, die derartig sicher wohl kaum eintreffen wird. Vertiefen Sie sich in Gedanken rasch in diese Situation. Notieren Sie in kurzen Stichworten, welche Folgen es haben würde und was alles passieren könnte, wenn diese Situation tatsächlich eingetreten wäre.

Verzichten Sie auf Begründungen und umständliche Erklärungen. Schreiben Sie Ihre Einfälle im Telegrammstil einzeilig untereinander auf. (Schoppe, 1975, S. 8, zitiert nach Acker, 2001, S. 80)

Laut Schoppe misst dieser Subtest den „Abruf von Einfällen, bei denen die logische Exploration von Erfahrungen in neuen Situationen gefordert ist.“ (Acker, 2001, S. 80)

Item 1:

„Welche Folgen würde es haben, was könnte alles passieren, wenn es für einen Monat überhaupt keinen Strom mehr gäbe?“

2.2.11. Wiedererkennen (WEK)

Zu dieser Dimension zählen Verfahren, bei denen Testpersonen versuchen müssen, aus einer Reihe vorgegebener Wörter das einzig wirklich reale Wort zu finden. Diese Aufgabentypen messen vorwiegend die Wiedererkennungslleistung. Ein Verfahren, das diese Dimension erfasst ist der WST.

Beispiel aus dem WST:

Der Wortschatztest von Schmidt & Metzler (1992) besteht insgesamt aus 42 Items, wobei jedes davon aus sechs „Wortgebilden“ zusammengesetzt ist. Die Testperson muss nun versuchen das richtige Wort zu identifizieren.

8. Malek - Kelmak - Almek - Makel - Nastel - Akelm

2.2.12. Wortverständnis (WV)

Auch dieser Dimension kann wieder nur ein Testverfahren, nämlich der VKI, zugeordnet werden. Hier geht es kurz gesagt um die Zuordnung vorgegebener Wörter zu inhaltlich entsprechenden Bildern.

Beispiel aus dem VKI:

Der Verbale Kurz-Intelligenztest von Anger et al. (1980) soll laut Autoren vor allem die verbale Intelligenz und die allgemeine Urteils- und Denkfähigkeit messen.

Jede Form enthält 20 Wörter und 4 einfache Darstellungen alltäglicher, leicht verständlicher Situationen. Der Proband hat die Aufgabe, zu jedem Wort dasjenige Bild zu suchen, das am besten dazu paßt. Wörter und Zeichnungen sind so ausgewählt, daß in jedem Fall eindeutige Optimallösungen erfolgen können (Anger et al 1980, S. 3, zitiert nach Acker, 2001, S. 83).

Item 15: „Mediation“

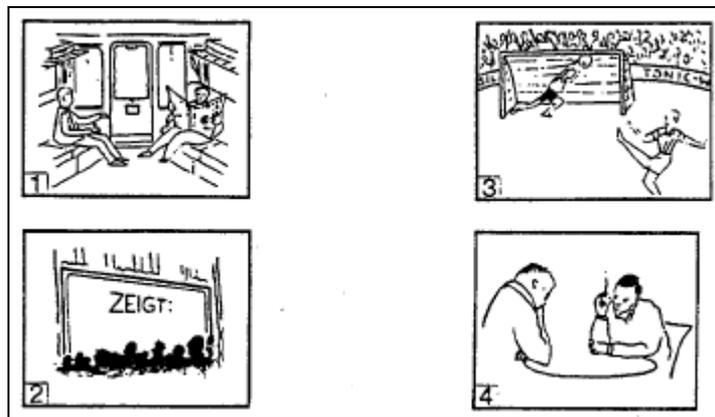


Abbildung 6: „Beispiel aus dem Verbalen Kurz-Intelligenztest (VKI) von Anger et al 1980“ (Acker, 2001, S. 84)

Aufgrund des Themas der vorliegenden Arbeit soll nun den zwei Dimensionen „Fluency“ und „Synonyme“ besondere Aufmerksamkeit geschenkt werden.

2.2.13. Synonyme (SYN)

Ein weit verbreiteter Aufgabentyp zur Messung der sprachlichen Intelligenz ist die Vorgabe von Synonymen. Tests, die hier genannte werden können, sind beispielsweise: AID Subtest 6 (Synonyme Finden), CFT 20 Ergänzungstest (Wortschatztest), HAWIE-R Subtest 5 Zusatztest (Wortschatztest – WT), KFT 4-13⁺ Subtest 1 (V1 Wortschatz), M-I-T Subtest 2 (Wortbedeutungen – WB), SASKA Subtest 1 (Sy 1-3), VKT Subtest 6 (Ähnlichkeiten – ÄK),

W-I-T Subtest 2 (Gleiche Wortbedeutung) (Acker, 2001). Als relativ neuer Test kann hier auch der LEWITE von Wagner-Menghin 2004 angeführt werden.

Adaptives Intelligenz Diagnostikum von Kubinger & Wurst (1985)

(Synonyme Finden – SF) - SYN^{akt}

Der Subtest 6 des AID versucht den aktiven Wortschatz einer Testperson zu prüfen. Die Aufgabe besteht darin zu den vorgegebenen Worten ein weiteres mit gleicher Bedeutung zu finden. Da keine schriftlichen Antwortalternativen vorgegeben werden, muss aus dem aktiven Wortschatz ein entsprechender Begriff gefunden werden. Das Verfahren ist ein reiner power-Test, der ohne Zeitdruck vorgegeben wird. (Acker, 2001)

Grundintelligenzskala 2.3. von Weiß (1987) - Ergänzungstest (Wortschatztest) – SYN^{pass}

Beim CFT-20 handelt es sich um einen passiven Wortschatztest, bei dem aus fünf verschiedenen Antwortalternativen ein bedeutungsähnliches Wort zu einem Distraktor gefunden werden soll. Aufgrund der Durchführung unter Zeitdruck kommt hier wieder die speed-power-Problematik zu tragen. Zum Inhalt der Items muss kritisch angemerkt werden, dass einige Begriffe zu Missverständnissen führen können, wie die Beispiele 11, 12, 25 und 28 zeigen und der Test insgesamt zu einfach gestaltet ist. (Acker, 2001)

| | | | |
|-----------------------|----------------------|----------------------|----------------------------|
| 11) Anwalt | 12) Miete | 25) Sehnsucht | 28) Droge |
| a) Scheidung | a) Wohnung | a) Liebe | a) <u>Betäubungsmittel</u> |
| b) Landschaft | b) Schwäche | b) Gewinn | b) Reitertruppe |
| c) Gehalt | c) Haus | c) <u>Wunsch</u> | c) Apotheke |
| d) Ansicht | d) <u>Geldbetrag</u> | d) Freude | d) Getreide |
| e) <u>Verteidiger</u> | e) Gewinn | e) Glück | e) Spritze |

Hamburg Wechsler Intelligenztest für Erwachsene Revision von Tewes (1991)

(Wortschatztest – WT) - SYN^{akt}

Der Subtest 5 des HAWIE-R listet 32 Wörter auf, zu denen selbstständig Wörter mit ähnlicher Bedeutung gefunden werden sollen. Bei diesem Verfahren muss allerdings auch unter Zeitdruck gearbeitet werden.

Die Leistung im Wortschatz-Test gilt nach Matarazzo als ‚exzellentes Maß‘ der allgemeinen Intelligenz einer Person (Matarazzo, 1982, S. 298), die auch weitgehend unabhängig vom Lebensalter ist. Die Leistung im Wortschatz-Test sei ein gutes Maß für die Lernfähigkeit und die verbale Informationsbreite des Probanden. Wortschatz-Tests gab es schon bei BINET (1905), TERMAN und CHAMBERLAIN (1981) und bei WHIPPLE (1921) (Tewes, 1991, S. 16, zitiert nach Acker, 2001, S. 72).

„Was bedeutet ...?“

1. Apfel
8. anonym
11. Naht
19. Deformation
32. Geoid

LEWITE - Lexikon-Wissen-Test von M. Wagner-Menghin (2004)

Dieses Verfahren geht auf die Theorie von Cattell zurück und versucht über die Erfassung des Wortschatzes die kristalline Intelligenz (gc) einer Testperson zu messen. Weiters realisiert dieser Test auch in Form eines objektiven Persönlichkeitstests, die Erfassung des wissensbezogenen Selbstkonzeptes. Die getestete Person wird zuerst aufgefordert einzuschätzen, ob Sie ein bestimmtes Wort erklären könnte. Erst danach erfolgt die Überprüfung anhand eines Lückentextes mit einem 2x4 Antwortformat, wobei die Testperson

eine sinnvolle Definition des zu erklärenden Begriffes rekonstruieren soll.

(www.schuhfried.at, Stand 10.9.2008)

Anleitung...

Versuchen Sie nun bitte den Satz zu vervollständigen, auch wenn Sie den Begriff und seine Bedeutung nicht kennen.

FRIES ist ein _____

gliedernder Schmuckstreifen beim Schneidern

inverser Rapport in der Baukunst

verbindendes Stück im Gartenbau

symmetrisch angelegtes Bauwerk

Zurück Weiter

Abbildung 7: Beispielitem aus dem LEWITE“ (<http://www.schuhfried.at>, Stand 10.9.2008)

2.2.14. Flüssigkeit (FLU)

In faktorenanalytischen Untersuchungen ist der Begriff „Fluency“ erstmals durch Hargreaves (1928) zum Untersuchungsgegenstand geworden. Die ersten Fluency-Tests wurden von Cattell 1935 konstruiert und mit Thurstones Modell der „Primary Mental Abilities“ erhielt der Faktor „Word-Fluency“ einen neuen Stellenwert innerhalb der Intelligenzstrukturforschung (Meili, 1981).

Heutzutage findet man Aufgaben zur Wortflüssigkeit oft in der neuropsychologischen Diagnostik zur Erfassung des divergenten Denkvermögens. Es kann als „flüssiges, originelles Denken, bei dem die Produktion möglichst vieler Lösungsmöglichkeiten angestrebt wird“ beschrieben werden. „Divergente Aufgabenstrukturen stellen offene Probleme dar, bei denen weder das Ziel noch die Lösungswege eindeutig definiert sind“ (König, 1986, zitiert nach Aschenbrenner, 2000, S.9). In einigen Untersuchungen konnte festgestellt werden, dass es nicht nur einen Fluency Faktor gibt, sondern einige spezielle Faktoren unterschieden werden

können: „word fluency“, „associational fluency“, ideational fluency“, „spontaneous flexibility“, „adaptive flexibility“, „redefinition“ usw. (Meili, 1981, S. 159)

Ganz allgemein kann der Faktor „Fluency“ in Aufgaben mit möglichst vielen Antwortproduktionen zu einer vorgegebenen Information festgestellt werden. Eine Testperson soll somit zu ein und derselben Information möglichst viele Antworten produzieren. (Meili, 1981)

Die Leistungen einer Person in verbalen Flüssigkeitsaufgaben sehen Aschenbrenner et al. (2000) abhängig von der Fähigkeit zur spontanen Produktion von Wörtern, d.h. der Existenz eines intakten lexikalischen und semantischen Wissens sowie die Funktionstüchtigkeit der Koordinations- und Abrufprozesse aus dem Gedächtnis.

Um die Variation der Antwortproduktionen zu erklären, untersuchte Meili (1981) drei Hypothesen

1. Die Leichtigkeit des Abrufes:

Er geht hier von der Annahme aus, dass die Anzahl der Antworten von der Leichtigkeit des Abrufes mancher Gedächtnisinhalte (Evokationsprozess) abhängen. Darunter versteht man die Zeit zwischen Präsentation eines Signalements und der gegebenen Antwort. Radziwanowski führte diesbezüglich eine Untersuchung durch, bei der diese Hypothese allerdings mit großer Wahrscheinlichkeit falsifiziert werden konnte.

2. Umfang des verfügbaren Wortschatzes:

Für Meili steht es außer Zweifel, dass die Anzahl der genannten Antworten von den im Wortspeicher verfügbaren Einträgen beeinflusst wird. Man würde wohl meinen, dass eine Testperson mit einem sehr umfangreichen Wortschatz wahrscheinlich mehr Antworten finden kann, als jene mit einem kleineren Wortschatz. Diese Hypothese ist jedoch schwer zu überprüfen, da der wahre Wortschatz einer Person nicht gemessen

werden kann. Radziwanowski gestaltet jedoch diesbezüglich ein Experiment, bei dem vor dem Fluency Test relevante Wörter aktiviert werden. Doch auch durch die vorherige Aktivierung konnte die Anzahl der Antworten nicht erhöht werden. Aus den Ergebnissen zu dieser Untersuchung kann daher geschlossen werden, dass die Leistung nicht unbedingt von der Verfügbarkeit möglicher Antworten abhängt.

3. Der Übergang von einer Antwort zur folgenden:

Aufgrund gestalttheoretischer Überlegungen stellte Meili (1981) die Hypothese auf, dass die Leistungen in Flüssigkeitsaufgaben umso besser sind, je schneller eine Person von der gegebenen Antwort loskommen kann und das Signalement wieder wirksam wird. Radziwanowski hat dazu ein einfaches Experiment durchgeführt, wobei die Versuchspersonen möglichst viele Dinge nennen sollten, die biegsam sind. Nach sechs Monaten wiederholte er den Versuch. Diesmal bekamen die Testpersonen Papierstreifen, auf denen ein Signalement geschrieben war. Auf jeden dieser Streifen durfte nur eine Antwort notiert werden. Anhand dieser Versuchsanordnung sollte das rasche Zurückkehren zur primären Aufgabe sichergestellt werden. Diese Untersuchung konnte signifikant höhere Leistungen der Testpersonen in der Versuchsanordnung mit Rückführung zum Signalement feststellen. Meili (1981) geht von der Annahme aus, dass die Leistungsmotivation hier eine große Rolle spielt.

Neuere Studien, wie beispielsweise die von Ruff et al. (1997) weisen auf einen positiven Einfluss von Aufmerksamkeit, verbalem Langzeitgedächtnis und Wortschatz auf Wortflüssigkeitsaufgaben hin (Okupski, 2006).

Es können folgende Aufgabenarten zur Messung verbaler Flüssigkeit unterschieden werden:

- Formallexikalische Wortflüssigkeitsaufgaben:

Gemeint sind hier Aufgaben, die eine möglichst große Anzahl an Wörtern mit bestimmter formaler Struktur erfordern (z.B.: mit bestimmten Anfangsbuchstaben).

- Kategorial-semantische Wortflüssigkeitsaufgaben:

Anhand der Vorgabe bestimmter Kategorien sollen möglichst viele Beispiele gefunden werden. (z.B. Vertreter für Kategorie Tiere) (Aschenbrenner et al., 2000)

- Buchstabenvertauschung (Anagramm):

Wie im zweiten Teil dieser Arbeit noch näher erläutert wird, kann die Wortflüssigkeit auch anhand der Vorgabe von Wörtern mit vertauschten Buchstaben gemessen werden. Den Testpersonen wird ein Wort vorgegeben, wobei die Buchstabenfolge durcheinander geraten ist. Die Aufgabe besteht nun darin, das Wort zu erkennen.

Gallistl (2006) beschreibt diesen Aufgabentyp als „Anagramm“. Im Historischen Wörterbuch der Rhetorik kann folgende Definition dafür gefunden werden:“ Unter Anagramm versteht man die Umstellung von Buchstaben innerhalb von Worten, Syntagmen, Sätzen und insbesondere von Namen, so dass eine andere Buchstabenfolge mit neuem Sinn entsteht... Metagramm, Kryptogramm oder Kryptonim sind andere Bezeichnungen für die Technik der Buchstabenversetzung.“ (Brunnschweiler, 2004, S.18)

Dieser Definition gemäß, kann der Aufgabentyp des hier verwendeten Tests nicht eindeutig den Anagrammen zugeordnet werden, da nicht durch Umstellung der Buchstaben eines wirklichen Wortes ein neues gefunden werden soll, sondern lediglich die durcheinander geratene Buchstabenreihenfolge zu einem Wort zusammengesetzt werden soll.

Controlled Oral Word Association Test von Benton & Hamsher (1976)

Im englischen Sprachraum ist vor allem der COWA das wohl am häufigsten angewendete Verfahren zur Messung der Wortflüssigkeit (Aschenbrenner et al., 2000).

Die Testpersonen sollen innerhalb einer Minute möglichst viele Wörter zu einem bestimmten Anfangsbuchstaben finden. Wörter mit gleichem Wortstamm, Wiederholungen, Namen oder Plätze werden als Fehler verrechnet. Der Rohwert wird anhand der richtigen Antworten der drei Aufgaben errechnet. Der Test differenziert nach Alter, Bildung und Geschlecht (Ruff et al., 1997, zitiert nach Okupski, 2006).

Regensburger Wortflüssigkeits-Test von Aschenbrenner, Tucha und Lange (2000)

Der aus 14 Untertests zusammengesetzte RWT ist ein Test zur Messung der verbalen Wortflüssigkeit. Das Verfahren unterscheidet formallexikalische Wortflüssigkeitsaufgaben und semantische Wortflüssigkeitsaufgaben. Je nach Untertest werden die Testpersonen gebeten innerhalb von zwei Minuten zu einem Anfangsbuchstaben bzw. zu einer vorgegebenen Kategorie so viele Wörter wie möglich zu nennen. Wörter mit gleichem Wortstamm oder Eigennamen werden als Wiederholungen verrechnet. Beide Untergruppen können weiter in unterschiedlich große „Suchräume“ (sehr groß, groß, mittel, gering) unterteilt werden. Der Suchraum kann als „Worthäufigkeit der Wörter, die mit dem vorgegebenen Buchstaben beginnen“ bzw. anhand der Bekanntheit der Kategorien definiert werden (Aschenbrenner et al, 2000, S.13). Nachdem die motorische Leistungsfähigkeit vieler Patienten eingeschränkt ist, wurde bei diesem Verfahren auf eine schriftliche Leistungsprüfung verzichtet (Aschenbrenner et al., 2000).

Beispiel:

- formallexikalische Wortflüssigkeit: Wörter die mit S beginnen
- semantisch-kategorielle Flüssigkeit: Tiere

Diagnostischer Test Deutsch von Nauck & Otte (1980)

(aktiver Wortschatz – AW) – FLU⁺ 2

Der sechste Subtest DTD setzt sich aus zwei Items zusammen, wobei jeweils zu einem vorgegebenen Buchstaben innerhalb von zwei Minuten möglichst viele Wörter produziert werden müssen. Laut Autoren ist die Leistung primär vom passiven Wortschatz abhängig.

| K,k | U,u |
|------------|------------|
| Kurz | unser |
| Kranz | Unke |
| ... | ... |

Acker (2001) weist darauf hin, dass die Schreibfähigkeit und natürlich auch die -geschwindigkeit maßgeblich an der Leistung bei diesen Aufgaben beteiligt sind.

Leistungsprüfsystem von Horn (1962)

Die Untertests 5 und 6 des LPS beinhalten insgesamt 80 Aufgaben zur Wortflüssigkeit. Im Subtest 5 ist die Aufgabe der Testpersonen, eine vorgegebene Buchstabenreihe so umzustellen, dass ein richtiges Wort gebildet werden kann. Hat die Testperson das Wort erkannt, muss sie den ersten Buchstaben des richtigen Wortes durchstreichen. Zur Bearbeitung dieses Subtest werden drei Minuten festgesetzt (Horn, 1962).

GZWER → ~~G~~ZWER (Zwerg)

² Mit ...⁺ gekennzeichnete und von Acker (2001) analysierte Tests weisen die speed-power-Problematik auf. Zudem beeinflusst die Schreibfähigkeit und-geschwindigkeit einer Testperson die Leistung. Quantitative Aspekte stehen im Vordergrund.

Prüfsystem für Schul- und Bildungsberatung von Horn (1969) - FLU⁺

Das PSB ist eine Kurzform des inzwischen ebenso veralteten Leistungsprüfsystems von Horn (1962). Der Test besteht insgesamt aus zehn Untertests, wobei sich zwei davon auf die verbalen Fähigkeiten konzentrieren. Aufgabe der Testperson ist es innerhalb von jeweils einer Minute möglichst viele Wörter zu einem vorgegeben Anfangsbuchstaben zu finden und aufzuschreiben. Kritik an dieser Aufgabenstellung ist wiederum das Problem der Schreibgeschwindigkeit, der Rechtschreibkenntnisse sowie die Schwierigkeit divergentes Denken unter Zeitdruck zu messen (Acker, 2001).

L,l - lernen, ...

P,p - Pech, ...

WILDE-Intelligenz-Test von Jäger & Althoff (1983)

(Wortgewandtheit – WG) – FLU⁺

Im Subtest 12 des W-I-T werden jeweils die Anfangs- und Endbuchstaben vorgegeben. Die Testperson soll nun innerhalb einer Minute so viele passende Wörter wie möglich finden. Durch den Zeitdruck kann auch hier wieder der kreative Denkprozess gestört werden. Die Autoren weisen darauf hin, dass die Untertests nicht eindimensional messen, wodurch der angegebene Verrechnungsmodus (Anzahl der gelösten Aufgaben = Maß für die Testleistung) nicht zulässig ist (Acker, 2001).

L – E

Beispielwort: Lage, ...

Verbaler Kreativitätstest von Schoppe (1975) - (Wortanfänge – WA) – FLU^{+ASS} ³

Der VKT Subtest 1 gibt Wortanfänge vor, zu denen so viele Wörter wie möglich gebildet werden sollen. Nachdem hier mehr als ein Buchstabe vorgegeben wird, kommt bei diesem Test dem assoziativen Aspekten eine größere Bedeutung zu (Acker, 2001).

(A) Item 1: Vorsilbe „**VER-**“ z.B.: verlieben, verfluchen, verzaubern, ...

Verbaler Kreativitätstest von Schoppe (1975) - (Wortenden – WE) - FLU^{+ASS}

Im Subtest 2 des VKT werden die Testpersonen nun angewiesen so viele Wörter wie möglich zu einem vorgegebenen Wortende zu finden und aufzuschreiben.

„Es können groß- oder kleingeschriebene Wörter jeder Art sein, also auch Fremdwörter, Namen, usw. [...], jedoch keine direkt fremdsprachlichen Wörter“ (Schoppe 1975, TH, S. 2, zitiert nach Acker, 2001, S.49). Nach Schoppe (1975) soll hier der „Abruf von Assoziationen nach vorgegebenen Ähnlichkeiten der Wortgestalt“ gemessen werden (S. 32, zitiert nach Acker, 2001, S. 50).

Item 1: „**-LOS**“ z.B.: bedeutungslos, schwerelos, ...

³ Die Kennzeichnung ...^{ASS} bezieht sich auf einen zusätzlichen assoziativen Aspekt

3. Mentales Lexikon

Was ist eigentlich das mentale Lexikon? Auch wenn die irreführende Bezeichnung als Lexikon eine mentale Repräsentation alphabetisch geordneter Worte nahe legt, so handelt es sich doch um eine komplex angelegte, systematische Struktur, die es uns ermöglicht, auf eine sehr effektive Art und Weise einen schnellen Zugriff auf die gespeicherten Einträge zu gewährleisten, auf die im folgenden Kapitel näher eingegangen werden soll.

Im Gegensatz zur Auffassung des Laien gilt der Wortschatz einer Person nicht nur als Index seiner schulischen Ausbildung, sondern auch als ausgezeichneter Maßstab seiner allgemeinen Intelligenz. Die Anzahl der Worte, über die ein Mensch verfügt, bildet ein Maß seiner Lernfähigkeit, seines Bestandes an sprachlichen Kenntnissen und seines allgemeinen Vorstellungsumfanges (Wechsler 1961, S. 113, zitiert nach Schwibbe, 1984, S. 15).

Glück (2005) beschreibt den menschlichen Wortspeicher als „mentale Organisation des Wortvorrates“, welche „interne Organisationsprinzipien, Anlage und Aktivierung von Worteinträgen, Zugang zu Wortrepräsentationen im Sprechen und Verstehen, deren Weiterverarbeitung, Verhältnis des mentalen Lexikons zum Sprach- und Weltwissen und zur mentalen Grammatik, Repräsentation des gesprochenen vs. des geschriebenen Wortes“ zum Gegenstand hat. (S.402)

Die meisten von uns verwenden ständig Worte, ohne weiter darüber nachzudenken, wie sie in unseren Köpfen gespeichert sind, noch wie wir darauf zugreifen können.

Nach Aitchison (1997) ist dies allerdings eine bemerkenswerte Leistung, wenn man sich vor Augen führt, dass ein erwachsener gebildeter Mensch mit Englisch als Muttersprache wahrscheinlich mehr als 50000 Wörter kennt und die meisten davon auch aktiv nutzen kann.

Anzumerken ist an dieser Stelle jedoch, dass diese Zahl nur bedingt der Realität entspricht, da die Definition eines „Wortes“ umstritten ist und sich die Vokabelkenntnisse nur schwer schätzen lassen. Die hier angegebene Schätzung bezieht sich auf Worte definiert als „im Wörterbuch eingetragen“. Es wird daher angenommen, dass sich die wahre Anzahl der gespeicherten Worte weit darüber befindet. Wenn man bedenkt, wie lange man braucht um ein Wort im Wörterbuch nachzuschlagen, dann erscheint es umso bemerkenswerter, dass ein Sprecher ein Wort in seiner Muttersprache innerhalb einer Fünftelsekunde erkennen kann (Aitichson, 1997).

Miller (1993) versucht diese Komplexität folgendermaßen auszudrücken:

Es sind nicht die Sprachlaute oder die Regeln zur Bildung grammatischer Sätze, die den größten Lernaufwand erfordern – es ist der Wortschatz: Tausende von Wörtern jedes mit seiner eigenen Aussprache, seiner eigenen Schreibweise, seiner eigenen Bedeutung, seiner eigenen Wortart, seiner eigenen Verwendung, seiner eigenen Geschichte (S. 18).

Jescheniak beschreibt, dass ein Sprecher normalerweise 2-3, manchmal aber auch bis zu 7 Worte pro Sekunde produzieren kann, wobei dies meist auch fehlerfrei geschieht.

Schätzungen zu Folge kommt nicht mehr als ein Wort- bzw. Lautfehler pro 1000 produzierter Wörter vor (Deese, 1984; Levet, 1989; Maclay & Osgood, 1959, zitiert nach Jescheniak, 2002).

Es scheint daher plausibel, dass die Wörter in unseren Köpfen sehr gut organisiert sein müssen, zumal unser Wortspeicher auch kein begrenztes Repertoire an Einträgen verwaltet. Es stellt vielmehr ein „lebendiges System“ dar, welches ständig seine Verbindungen erweitert (Aitichson 1997).

3.1. Methoden zur Erforschung der Struktur des mentalen Lexikons

Um den menschlichen Wortspeicher erforschen zu können, bedient sich die Wissenschaft unterschiedlicher Methoden, die Aufschluss über diese komplexe Struktur bringen sollen.

Nach Blutner (1986) ist eine unmittelbare Beobachtung der Vorgänge im mentalen Lexikon nicht möglich, weshalb auf indirekte Methoden zurückgegriffen werden muss. Verfahren, die zur Klärung herangezogen werden sind beispielsweise Experimente zur Wortidentifizierung, lexikalische Entscheidungsaufgaben sowie Nachsprechaufgaben.

Einige Methoden, die in Jescheniak (2002) dazu beschrieben werden, sind Analysen von Sprechfehlern, Wortfindungsstörung hirngesunder Menschen und von amnestischen Patienten, sowie chronometrische Studien.

3.1.1. Sprechfehler

„Der Zufall ist beim Versprechen vollkommen ausgeschlossen, das Versprechen ist geregelt“ (Rudolf Meringer, 1908, zitiert nach Jescheniak, 2002, S. 34).

Wie das Zitat schon auszudrücken versucht, vermutet man hinter jedem Versprecher eine gewisse Systematik. Durch diese Produktionsstörung erhofft man Erkenntnisse über unbewusste Prozesse gewinnen zu können.

Folgende Sprechfehler können unterschieden werden:

- Wortvertauschungsfehler (z.B. a wife for his job – a job for his wife)
- Lautvertauschungsfehler (z.B. heft lemisphere – left hemisphere)
- Wortersetzungsfehler
 - Mit semantischer Relation (z.B. don't burn your toes – don't burn your fingers)
 - Mit phonologischer Relation (z.B. bottle of page five – bottom of page five)

- Mit semantischer und phonologischer Relation (a book with a most magnificent dialect – a book with a most magnificent dialogue)
(Jescheniak, 2002)

3.1.2. Wortfindungsstörung bei hirngesunden Sprechern

Jeder von uns kennt den Zustand, wenn einem ein bestimmtes Wort “auf der Zunge liegt”. Reason (1984, zitiert nach Jescheniak, 2002) berichtet, dass ca. 51% der interviewten Personen wöchentlich so ein TOT-Erlebnis (Tip of the tongue) haben. Bei Untersuchungen zu diesem Phänomen konnte festgestellt werden, dass in 51 – 71% der Fälle der erste Buchstabe richtig wiedergegeben werden konnte. In manchen Fällen können wir uns auch noch an den letzten Buchstaben erinnern.

Dieses Phänomen beschreibt Aitchison (1997) als den „Badewanneneffekt“, der dadurch charakterisiert ist, dass genauso wie in der Badewanne, der Kopf (= der initiale Buchstabe) weiter aus dem Wasser ragt als die Füße (=Endbuchstabe), der Rest ist unter dem Wasser. Auch Informationen bezüglich der Anzahl der Silben konnten oft abgerufen werden. Einige Autoren gehen von der Annahme aus, dass eine ungenügende Aktivierung der Wortform verantwortlich sei (Levet, 1989; Burke et al, 1991, zitiert nach Jescheniak, 2002). Eine weitere Erklärungsmöglichkeit bietet Jones und Langford 1987 (zitiert nach Jescheniak, 2002), indem ähnlich klingende Wörter, die beim Suchprozess in den Sinn kommen, den Zugriff auf das Zielwort erschweren.

3.1.3. Wortfindungsstörung bei amnestischer Aphasie

Nicht nur Sprachfehler hirngesunder Personen können Aufschluss geben, sondern auch Störungsmuster aphasischer Patienten. Typisch für diesen Defekt ist eine gravierende

Wortfindungsstörung, einem anomischen Zustand, der mit intaktem Sprachverstehen einhergeht. (Jescheniak, 2002)

3.1.4. Chronometrische Studien

Anhand dieser Studien sollen die zeitlichen Abläufe der verschiedenen Teilprozesse der Lexikalisierung beschrieben werden. Ein Beispiel dafür ist das „Bild-Wort-Interferenzparadigma“, anhand dessen eben diese Aspekte näher beleuchtet werden sollen. Erstmals wurde es von Schiefer et al 1990 für diese Zwecke eingesetzt. Die Versuchspersonen sollen einfache Bilder benennen (z.B. ein Schaf), während ein akustischer beziehungsweise visueller Ablenkreiz, den sie ignorieren sollten, dargeboten wird. Visuelle Ablenker, die dem Zielwort ähnlich sind, verlangsamten die Reaktionszeit, während phonologisch ähnliche Reize den Zugriff erleichtern. (Jescheniak, 2002)

Eine ähnliche Untersuchung startete Meyer und Schvaneveldt (1971, zitiert nach Engelkamp, 1995) Anfang der 70er Jahre, in der sie zeigten, dass bedeutungsähnliche Worte sich gegenseitig beeinflussen. Den Personen wurden Wörter und Nichtwörter vorgegeben, die sie identifizieren mussten. Teils wurden davor bedeutungsähnliche Worte dargeboten und teils lautähnliche Reize. Es konnte hier eine verkürzte Reaktionszeit bei ähnlich phonologischen und semantischen Wörtern festgestellt werden. Diese beiden Effekte werden als „phonologischer“ und „semantischer Priming-Effekt“ bezeichnet. Nach Forster (1987, zitiert nach Engelkamp, 1995) kann der phonologische Priming-Effekt nur dann festgestellt werden, wenn das zuvor genannte Wort noch nicht erkannt wurde. Diese unterschiedlichen Priming-Effekte beschreibt Engelkamp (1995) als Evidenz für getrennte Netzwerke für Konzepte und Wortrepräsentationen.

3.2. Die Bedeutung von Wörtern

Um eine Sprache korrekt anwenden und verstehen zu können, ist es notwendig die Bedeutung der Wörter zu verstehen. Wie kann die Bedeutung allerdings definiert werden. Zwei Versuche sehen wir bei Herkner (2001): „Aus lernpsychologischer Sicht ist es naheliegend, die Bedeutung eines Wortes als dessen <richtigen> Gebrauch aufzufassen“ (S.141) und „Die Bedeutung eines Wortes ist die Menge seiner *semantischen Komponenten*“ (S. 144). Wie wir diese Bedeutungen erfassen, sollen nun folgende Theorien zu erklären versuchen.

3.2.1. Die Checklistentheorie

Aitichson (1997) beschreibt, dass einige Philosophen von der Annahme ausgehen, dass eine „Menge notwendiger und hinreichender Bedingungen“ vonnöten sind, um die Bedeutung eines Wortes zu umreißen. So führte er in diesem Zusammenhang das Beispielwort „Quadrat“ an. Folgenden vier Voraussetzungen sind dazu notwendig:

- Geschlossene, flache Figur
- Vier Seiten
- Alle Seiten müssen gleich lang sein
- Alle Innenwinkel müssen gleich sein

Dieser Auffassung nach bestehe im Gehirn des Menschen für jedes Wort eine Reihe von Voraussetzungen, eine so genannte Checkliste, die man unbewusst prüft, bevor man eine bestimmte Bedeutung zuordnet. Die Auswahl der notwendigen Kriterien stellt uns hier allerdings vor schwerwiegende Probleme. So hat sich gezeigt, dass sich für einige Dinge keine wesentlichen Merkmale erstellen lassen, da es für manche Worte keine festen Bedeutungen gibt. Als Beispiel dafür nennt Aitichson (1997) das Experiment des Soziolinguisten William Labov. Er zeigte seinen Studenten Abbildungen von verschiedenen Gefäßen, die sie als Tasse, Vase oder Schale klassifizieren sollten. Es konnte gezeigt werden,

dass hohe, schmale Gefäße ohne Henkel eindeutig als Vasen identifiziert werden konnten, während niedrige flache als Schalen bezeichnet wurden. Schwierig wurde es allerdings dann, wenn man ihnen ein „Zwischending“ gezeigt hat. Einen weiteren Kritikpunkt an der Checklistentheorie stellt das Syndrom der Familienähnlichkeiten dar. Zum Beispiel weisen manche Spiele Ähnlichkeiten auf, andere hingegen haben wenig gemeinsam. Trotzdem können sie als Spiel identifiziert werden, ohne unbedingt gemeinsame Merkmale aufweisen zu können. Diese Erkenntnis legt nahe, dass es wohl einige Begriffe gibt, die feste Kernbedeutungen besitzen, der Großteil jedoch weist unscharfe Grenzen auf.

3.2.2. Prototypentheorie

Diese Theorie postuliert, dass beim Klassifizieren von Objekten ein Vergleichsvorgang mit einem gewissen „Prototypen“ stattfindet. So geht man davon aus, dass bestimmte Vertreter einer Kategorie typischer sind als andere. Beispielsweise ist ein Rotkehlchen ein besserer Vertreter der Kategorie Vögel, als ein Strauß oder ein Pinguin. Diese Theorie ist der Checklistentheorie sehr ähnlich, allerdings ist es hier nicht eine bestimmte Anzahl an Merkmalen, die ein Objekt besitzen muss, um einer Kategorie zugeordnet werden zu können. Das Objekt muss den Prototypen nicht genau entsprechen, es müssen lediglich Ähnlichkeiten festgestellt werden können. Ein fluguntaugliches, einbeiniges Rotkehlchen kann somit ebenfalls als Vogel identifiziert werden. (Aitchison 1997)

3.3. Die Verbindung zwischen den Wörtern

Im folgenden Abschnitt soll die Frage erörtert werden, wie die Worte untereinander verbunden sind. Aitchison geht von der Annahme eines semantischen Netzwerkes aus, bei dem „die lexikalischen Einheiten die Knoten bilden, welche durch Pfade miteinander

verbunden sind.“(Aitchison 1997, S. 105). Unumstritten kann behauptet werden, dass irgendeine Art von Verbindung zwischen den Worten bestehen muss, über die Art dieser Verbindungen bestehen allerdings geteilte Meinungen. Anhand von Wortassoziationstests versuchte man die Verbindungen zu anderen Worten zu identifizieren.

Folgende Assoziationstypen können demnach unterschieden werden:

- **KOORDINATION:** Gemeint sind hier eine Gruppe von Wörtern, die auf der gleichen Detaillierungsebene liegen (z.B. Ausgangswort Rot: Schwarz, Gelb, usw.)
- **KOLLOKATION:** Wörter die in einem Kontext häufig gemeinsam auftreten (z.B. Salz: Salzwasser)
- **ÜBERORDNUNG:** Ein Oberbegriff oder auch „Hyperonym“, der das Ausgangswort umfasst (z.B. Rot: Farbe, Schmetterling: Insekt)
- **SYNONYME:** Ein Wort mit der selben Bedeutung (z.B. Leopard: Panther)

Eine vollkommene Bedeutungsgleichheit ist sehr selten, da nicht jedes Wort im gleichen Kontext verwendet werden kann. Es könnte beispielsweise ein Plan verfolgt werden, aber man kann ihn nicht jagen. Man nimmt an, dass Wörter mit einer ähnlichen Bedeutung in unterschiedlichen semantischen Bereichen gespeichert werden, da diese Art von Versprecher nur sehr selten auftreten. Hingegen Wörter, die gegeneinander ausgetauscht werden können, stehen vermutlich in enger Verbindung. Es kann nun angenommen werden, dass Wörter in sogenannten „semantischen Feldern“ angeordnet sind, innerhalb derer starke und schwache Verbindungen bestehen. Diese starken Verbindungen sollen nicht im Kontext von örtlicher Nähe verstanden werden, vielmehr soll der Aspekt der direkten, starken Aktivierung im Vordergrund stehen.

Weiters lässt sich feststellen, dass innerhalb der einzelnen Wortarten (Nomen, Verben und Adjektive) ebenfalls ausgeprägte Verbindungen bestehen, wobei die Anzahl der Wortklassen,

die im mentalen Lexikon unterschieden werden, nicht exakt festgestellt werden kann. Ein Sprecher hat somit schnellen Zugriff auf verwandte Worte und es erleichtert somit die Sprachproduktion. Aitichson postuliert, dass ebenfalls eine Unterscheidung in Funktionswörter und Inhaltswörter, die das eigentliche mentale Lexikon darstellen, getroffen wird. Die Bedeutung der Wörter und die Wortart scheinen gemeinsam gespeichert zu sein, sie bilden eine Einheit, die in der Literatur als Lemma bezeichnet wird. Aitichson beschreibt dies auf eine sehr anschauliche Art und Weise mit folgenden Worten „Wörter sind wie Münzen; Bedeutung und Wortart bilden gemeinsam die eine Seite (diese Einheit wird auch als *Lemma* bezeichnet), und die Lautgestalt die andere“ (Aitichson, 1997, S.128).

Wörter mit ähnlichem Klang dürften eine weitere Dimension der Verbindungen darstellen. Vor allem Analysen von Wortversprechern belegen die Annahme, dass aufgrund der Häufigkeit der Sprechfehler besonders starke Verbindungen zwischen sehr lautähnlichen Worten bestehen und die Verbindungen immer schwächer werden, je unähnlicher ein Wort klingt. Für die akustische Worterkennung stellt dies einen großen Vorteil dar. Oftmals können nicht alle Teile eines Wortes einwandfrei verstanden werden, so erscheint eine Menge von ähnlich klingenden Wörtern und es wird jenes ausgewählt, das am besten zu passen scheint. Diese Assoziationen können aber nicht nur helfen den Zugriff zu einem bestimmten Wort zu erleichtern, sie können in manchen Situationen regelrechte Blockaden darstellen, wie im TOT-Zustand schon erläutert wurde (Aitichson, 1997).

3.4. Zugriff auf das mentale Lexikon

Der menschliche Wortspeicher muss, um optimal arbeiten zu können, in der Lage sein, Informationen zu Syntax, Semantik und Phonologie parallel bearbeiten können. Auf den ersten Blick scheint es so, als wären Wortproduktion und Worterkennung zwei gegenläufige Prozesse. (Aitichson, 1997)

Im folgenden Absatz soll nun dieser Aspekt näher beschrieben werden.

3.4.1. Die Wortproduktion

Aitichson (1997) stellt dazu 3 Modelle vor, die Aufschluss darüber geben sollen, wie bei der Sprachproduktion die Wörter ausgewählt werden können. Wie man sieht, können die ersten beiden Theorien noch nicht alle Fragestellungen zur Genüge klären. Zwecks Vollständigkeit und zur besseren Veranschaulichung sollen nun auch diese beiden Modelle vorgestellt werden.

A) „*Das Trittsteinmodell*“ (Aitichson, 1997)

Dieses Modell versucht folgende Erkenntnisse zu berücksichtigen:

- Die Wortform und das Lemma sind getrennt gespeichert, dennoch bestehen gewisse Verbindungen.
- Versprecher sind oftmals dadurch gekennzeichnet, dass manche Teile des Zielwortes falsch gesprochen werden, während andere richtig formuliert werden.
- Bei der Auswahl der Worte stehen mehrere Alternativen zur Verfügung.

Das Trittsteinmodell nimmt an, dass eine Phase beendet sein muss, um mit der nächsten beginnen zu können. Der Informationsfluss verläuft hier nur in eine Richtung. Aitichson verdeutlicht diese Theorie anhand des Beispiels „Espe“. Zunächst werden auf dem ersten Trittstein die Bedeutung des Wortes und ihre Wortart ausgewählt, das als Einheit die Bezeichnung Lemma trägt. In unserem Beispiel weiß der Sprecher, dass es sich um eine bestimmte Art von Laubbaum handelt (= Bedeutung), sowie dass es sich um ein Nomen handeln muss (=Wortart). Nachdem auf dem ersten Trittstein der Prozess abgeschlossen ist, kann der nächste Schritt folgen. Die nächste Phase besteht nun darin die Lautung (Wortform)

zu eruieren. Damit im zweiten Schritt die Richtung der Suche festgelegt werden kann, muss bereits am ersten „Trittstein“ ein „Wegweiser“ angebracht werden, der in unserem Beispiel markiert, dass der gewünschte Ausdruck mit „E“ beginnt, mit „e“ endet und aus zwei Silben besteht. Allerdings sollte auch die Erkenntnis berücksichtigt werden, dass beim Suchen eines Wortes immer eine ganze Reihe von Worten mitaktiviert werden. Daher scheinen die „Trittsteine“ zu Beginn breiter angelegt zu sein, da noch mehr Möglichkeiten zur Auswahl stehen, aber beim Selektionsprozess immer enger werden je näher man dem Zielwort kommt. Auf der Lemmaebene werden beispielsweise eine Reihe von Nomen mit ähnlicher Bedeutung zu unserem Beispielwort „Espe“ erscheinen (Birke, Ulme, Esche, usw.). Hingegen werden auf der Lautungsebene klangähnliche Worte aktiviert (Ente, Erbe, Ende, Esche, usw.).

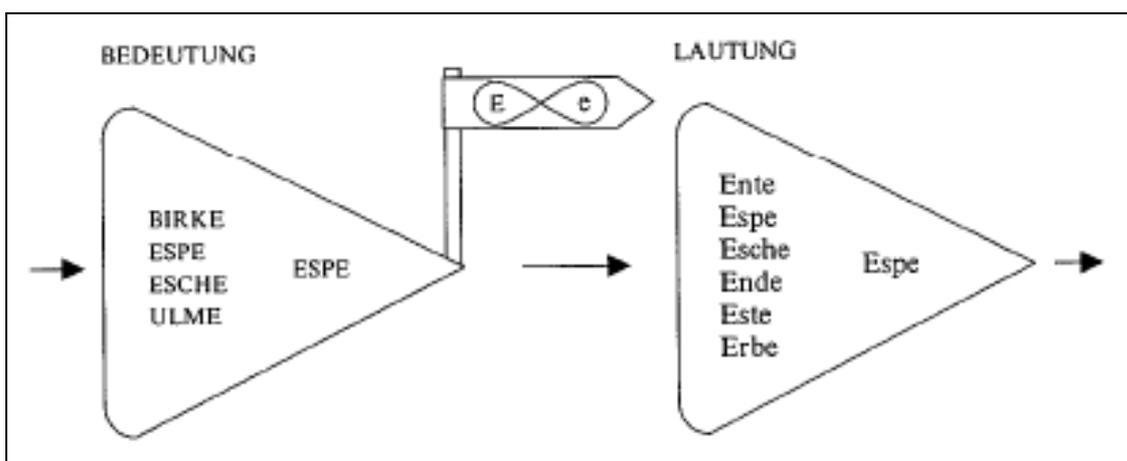


Abbildung 8: „Ein Trittsteinmodell mit Engpaß“ (Aitichson, 1997, S. 265)

Anhand dieses Modells ist es zwar möglich zu erklären, wie es auf den unterschiedlichen Ebenen zu falschen Selektionen kommen kann, allerdings erklärt es nicht Versprecher, die ähnliche Lautung und Bedeutung haben, da jede Phase abgeschlossen sein muss, bevor die darauf folgende beginnt (Aitichson, 1997).

Jescheniak (2002, zitiert nach Levet, 1989, 1992; Levet et al., 1991a, 1999a,b) nennt diese Theorie „Diskret-serielles Modell“. Nachdem diese Art von Fehler wahrscheinlich am

häufigsten auftritt, muss diese Theorie verworfen werden und es soll ein Modell beschrieben werden, das diese Erkenntnisse berücksichtigt.

B) Das „Wasserfallmodell“ (Aitichson, 1997)

Aufgrund der Probleme in der oben beschriebenen Theorie, war man bemüht ein Modell zu definieren, das eine überlappende Aktivierung der Phasen zulässt. Bei diesem Modell hat man somit noch Zugriff auf Informationen der vorigen Ebene, das heißt bevor noch eine Entscheidung auf der Bedeutungsebene gefallen ist, beginnt die Aktivierung der Lautstruktur.

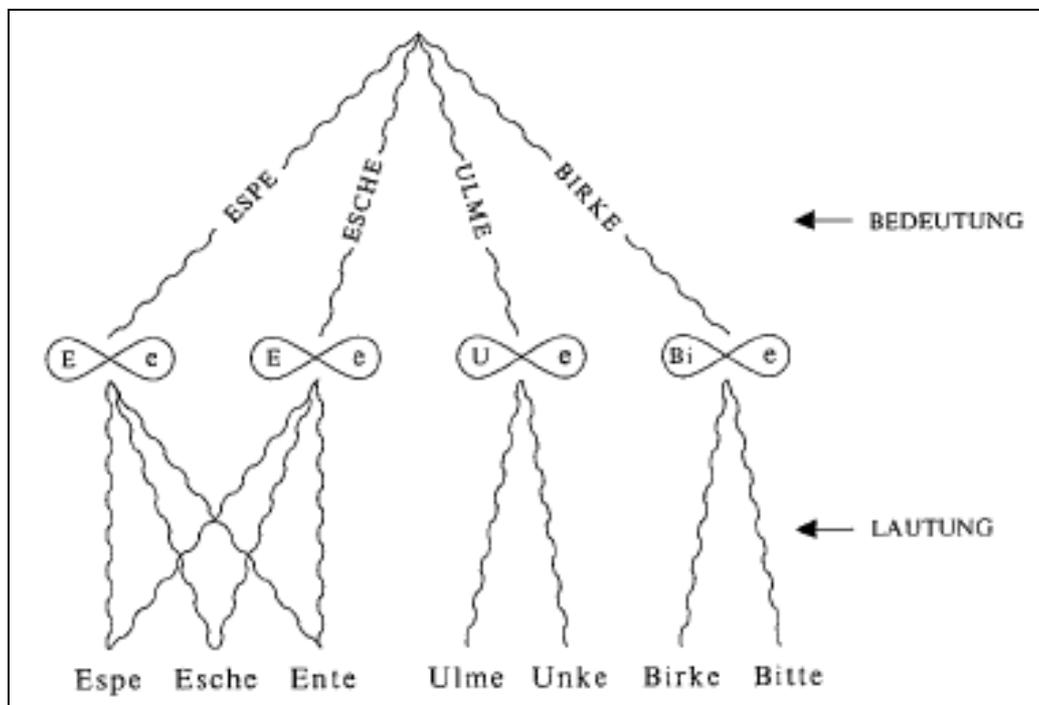


Abbildung 9: „Ein Wasserfallmodell“ (Aitichson, 1997, S. 267)

Leider weist auch diese Theorie noch einen gravierenden Mangel auf, da hier ebenfalls der Informationsfluss nur eine Richtung aufweist. Um die Lautstruktur finden zu können, braucht man nicht nur die Bedeutung, auch die umgekehrte Richtung ist oft wesentlich zum Finden eines Wortes. Stellt man jemanden die Aufgabe so viele Frühlingsblumen wie möglich zu nennen, erfolgen vielleicht Äußerungen wie: Schneeglöckchen, Narzisse usw. Erteilt man

jedoch den Hinweis auf eine Blume mit dem Anfangsbuchstaben T, so erinnert er sich vielleicht noch an Tulpe oder Traubenhyazinthe. Dieses Beispiel zeigt deutlich, dass die Informationen auch in die andere Richtung fließen können (Aitichson, 1997).

In der Literatur findet man dieses Modell auch noch unter dem Namen „vorwärts-kaskadierendes Modell“ (Cutting & Ferreira, 1998; Peterson & Savoy, 1998, zitiert nach Jescheniak, 2002)

C) Interaktives Aktivierungs-Modell (Aitichson, 1997)

Das „Wasserfallmodell“ oder „vorwärts-kaskadierende Modell“ berücksichtigt schon viele wesentliche Komponenten, doch soll hier noch versucht werden, der beidseitigen Ausrichtung des Informationsflusses Rechnung zu tragen. Als Metapher bietet sich hier ein unter Strom stehendes Netzwerk an, wobei der Strom zwischen den Knoten hin und her fließen kann. Bei der Suche des Wortes „Espe“ wird zunächst ein semantisches Feld (z.B. heimische Laubbäume) aktiviert. Bevor die Wahl noch auf ein bestimmtes Wort fällt, wird über „phonologische Vorwahlnummern“ zum Beispiel Anfangs- und Endbuchstaben eine Gruppe geeigneter Lautungen angeregt. Diese Aktivierung fließt nun wieder zurück zu den Bedeutungen und breitet sich somit aus. Es kommt zur Überprüfung der Aktivierungsknoten, wobei geeignete verstärkt werden, während nicht relevante immer schwächer werden. Nachdem starke Aktivierungen in der Semantik auch starke Erregungen im phonologischen Bereich auslösen und dies in entgegengesetzter Richtung ebenso funktioniert, kann es passieren, dass eine geeignete Lautstruktur zu mehreren Bäumen passt, so werden diese Einträge auch auf der Bedeutungsebene stärker aktiviert. Bei unserem Beispiel kann es nun vorkommen, dass bei Unachtsamkeit des Sprechers statt „Espe“ das Wort „Esche“ ausgesprochen wird.

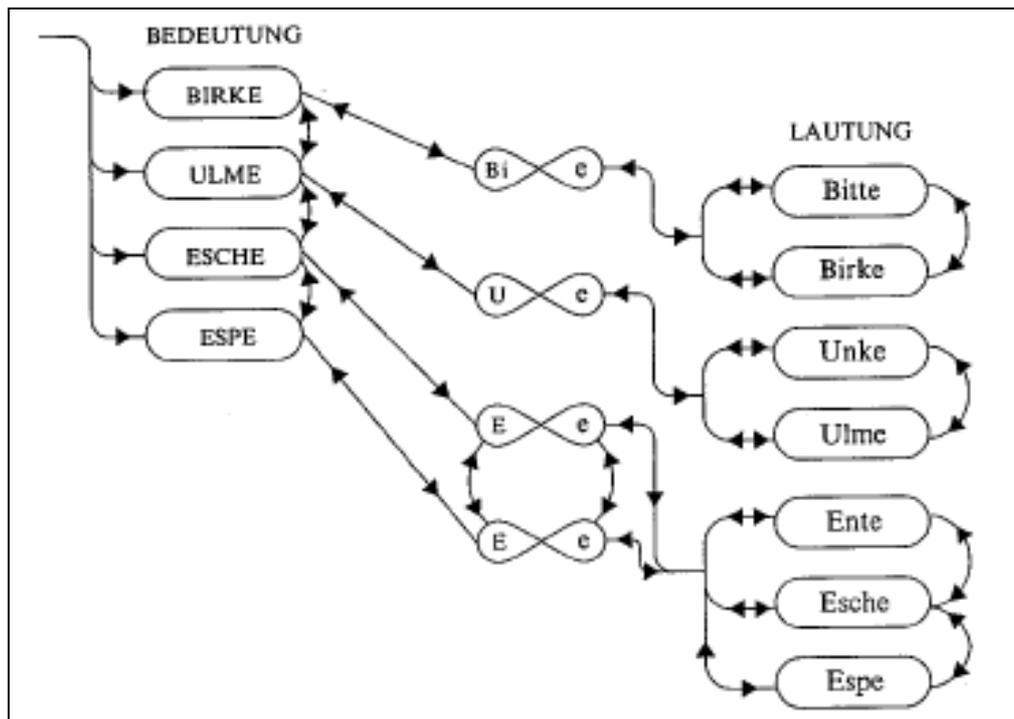


Abbildung 10: „Ein Interaktives Aktivierungs-Modell“ (Aitichson, 1997, S.269)

Es stellt sich nun die Frage, wann diese Aktivierungen zur Auswahl eines Wortes führen. Vielleicht muss die Aktivierung eines Wortes eine bestimmte Intensität erreichen, oder im Gegensatz zu den anderen einfach deutlich stärker erregt worden sein. Es wurde festgestellt, dass häufig verwendete Wörter sehr viel leichter zu aktivieren sind, als weniger gebräuchlichere. Oft kommt es vor, dass man schon kurz nach einem Versprecher entdeckt, dass man das falsche Wort ausgewählt hat. Deshalb ist es denkbar, dass nach der Auswahl eines Wortes eine „Kontrolleinrichtung“ nochmals überprüft, ob der Eintrag korrekt ausgesucht wurde (Aitichson, 1997).

3.4.2. Das Erkennen akustischer Wörter

Wie ich in der Einleitung schon erwähnt habe, ist ein Sprecher in der Lage 2-3 manchmal allerdings auch bis zu 7 Worte pro Sekunde zu produzieren. Oftmals führen Umgebungsgeräusche dazu, dass man Wörter akustisch nicht gut wahrnehmen kann, oder der

Sprecher verschluckt manche Buchstaben. Trotzdem gelingt es uns häufig die Worte einwandfrei zu erkennen. Aitichson (1997) beschreibt in seinem Buch, dass die Basis dafür oft lediglich Vermutungen sind. Der Zuhörer durchsucht sein mentales Lexikon nach dem Wort, das wohl am passendsten erscheint und füllt die Leerstelle, ohne es überhaupt bemerkt zu haben. Nun schauen wir uns die Prozesse im Einzelnen an.

Zu Beginn muss der Hörer versuchen den Redeschwall in einzelne Worte zu zerlegen, um dann die einzelnen Einheiten identifizieren zu können. Diese zwei Aufgaben muss unser Gehirn gleichzeitig bewältigen, wobei die Separierung der Worte einfacher zu gelingen scheint. Grundsätzlich kann zwischen einem „seriellen“ und einem „parallelen“ Modell unterschieden werden.

Das erstere besagt, dass jedes Wort nach dem anderen überprüft wird, während im zweiten dieser Vorgang gleichzeitig vonstatten geht. Die Theorie der schrittweisen Testung stützt sich vor allem auf die Tatsache, dass gebräuchlichere Wörter schneller identifiziert werden, als diejenigen, die wir nicht so oft benutzen. Oftmals sind dies Wörter, die wir schon sehr früh gelernt haben und die sich sprichwörtlich tiefer in unser Gedächtnis eingebrannt haben. Einige Wissenschaftler gehen von der Annahme aus, dass die Wörter gestapelt sind und die, die häufig Verwendung finden, oben aufliegen. Wenn wir ein Wort hören, so beginnen wir unsere Suche je nach Lautfolge in einer bestimmten Schublade, die wir von oben nach unten durchforsten. Nach Forster (1976, zitiert nach Aitichson, 1997) sind diese „Schubladen“ nach ihren Anfangsbuchstaben geordnet. Glanzer und Ehrenreich (1979, zitiert nach Aitichson, 1997) sehen dies etwas anders, demnach sind häufig gebrauchte Wörter doppelt gespeichert, einmal in einem Speicher mit leichtem Zugang und einmal an ihrem eigentlichen Ort.

Für das Modell der parallelen Bearbeitung sprechen vor allem die Daten von Untersuchungen zu Mehrdeutigkeiten. Auch die Tatsache, dass Wörter innerhalb kürzester Zeit erkannt werden können, legt nahe, dass der Zuhörer blitzschnell einige mögliche Alternativen zur Verfügung hat, von denen er im Anschluss aufgrund syntaktischer und semantischer Hinweise das

geeignetste auswählt. Das Kohortenmodell besagt, dass sobald der Sprecher beginnt Worte zu formulieren, eine „Wortarmee“, aufgrund der initialen Buchstaben ausgewählt wird und anhand des Kontextes zur Bevorzugung eines Wortes kommt. (Aitichson, 1997)

Blutner (1986) beschreibt, dass 150 – 200 ms nach dem Beginn der Artikulation schon die erste Kandidatenmenge gebildet wird.

Allerdings kann diese Theorie nicht erklären, wie man Worte richtig wahrnehmen kann, auch wenn die ersten Buchstaben falsch verstanden wurden. Aus diesem Grund wurde auch hier ein „Stromkreismodell“ formuliert, welches besagt, dass Lautform und Bedeutung sich gegenseitig aktivieren und hemmen können. Neu ankommende Informationen grenzen die Wortwahl immer mehr ein (Aitichson, 1997).

Blutner (1986) spricht in diesem Zusammenhang von dem „Einzigkeitspunkt“, an dem letztendlich ein Wort ausgewählt wird.

Somit können wir zusammenfassend feststellen, dass die Prozesse der Worterzeugung und der Worterkennung sehr ähnlich funktionieren. Der Unterschied besteht hauptsächlich im Ausgangspunkt, der bei der Produktion in der Bedeutung der Worte zu finden ist, während es bei der Erkennung die Laute sind, die den Anstoß geben. (Aitichson, 1997)

3.5. Die Gesamtstruktur des mentalen Lexikons

Um sich die Organisation des menschlichen Wortspeichers besser vorstellen zu können, versucht Aitichson (1997) eine Art „Landkarte“ zu erstellen. Er geht dabei von zwei großen Einheiten aus, die er als „Semstadt“ und „Phonstadt“ bezeichnet. In der ersteren sind Informationen zur Bedeutung und Wortart angesiedelt, während in der zweiten Angaben zu den Lauten (Wortformen) enthalten sind. Von diesen zwei „Städten“ ausgehend gibt es Verbindungen zu „Neustadt“, wo neue Wörter zusammengestellt werden können, wie

beispielsweise durch Kompositabildung⁴, Konversion⁵ und Affigierung⁶. Diese Städte liegen nur teilweise an der Oberfläche, ihre größten Teile erstrecken sich über riesige unterirdische Netzwerke. Diese großflächig angelegten Tunnel breiten sich nicht nur in die Tiefe aus, sondern haben auch seitwärts Verbindungen zu den anderen Städten. Manche dieser Tunnel sind leicht zugänglich und hell ausgeleuchtet, während andere noch erforscht werden müssen. Die Tunnel von „Semstadt“ haben weite Verbindungen zum Allgemeinwissen und zu Erinnerungen. Es ist daher schwierig festzustellen, wo „die Bedeutung eines Wortes endet und das Allgemeinwissen beginnt“ (Aitichson, 1997, S. 295). Weiters gibt es Verbindungen von „Semstadt“ zu allgemeinen syntaktischen Regeln. Der „Nebenspeicher“, der Wörter in Morpheme⁷ zerlegt, ist von allen Einheiten zugänglich. Es entspricht einer Art „lexikalischem Werkzeugkasten“, und befähigt uns, neue Wörter zu kreieren. Bei der Wortsuche gelangt man somit von den Eingängen in ein weit verzweigtes System von Netzwerken. Bei einem oft benutzten Wort sind die Pfade sozusagen schon breitgetreten und ein schnelles Vorankommen ist gesichert. Während der Weg bei seltenen Wörtern noch freigeekämpft werden muss. Jedes Wort hat Verzweigungen zu anderen Wörtern, sowie einen Zugang zu Einträgen aus dem allgemeinen Gedächtnis (Aitichson, 1997).

Das Bild der >Städte< und >Tunnel< ist nützlich, weil es den multidimensionalen Charakter der Verbindungen unterstreicht. Es verdeutlicht auch, dass sich wie bei wirklichen Städten, schwer bestimmen läßt, wo eine Stadt aufhört und die nächste beginnt. Es ist leichter, sie anhand ihres Zentrums zu identifizieren als anhand ihrer Grenzen“ (Aitichson, 1997, S. 296).

⁴ Kombination von Nomen z.B. Feuerwehrmann, Indianerhäuptling, usw. (Aitichson, 1997)

⁵ Hierbei wird eine Wortart in eine andere umgewandelt z.B.: „Den Kuchen mit Zucker bestreuen“ wird zu „den Kuchen *zuckern*“. (Aitichson, 1997)

⁶ Anhängen eines Präfix (z.B. *ent*-kalken) beziehungsweise eines Suffix (z.B. *Donald-Duckhaftigkeit*). (Aitichson, 1997)

⁷ Morphem = die kleinste selbstständig bedeutungstragende Einheit eines Wortes (Glück, 2005, S. 423) Z.B.: Das Wort „Personen“ besteht aus zwei Morphemen, nämlich „Person“ = freies Morphem und der Pluralendung „en“ = gebundenes Morphem. (Herkner, 2001)

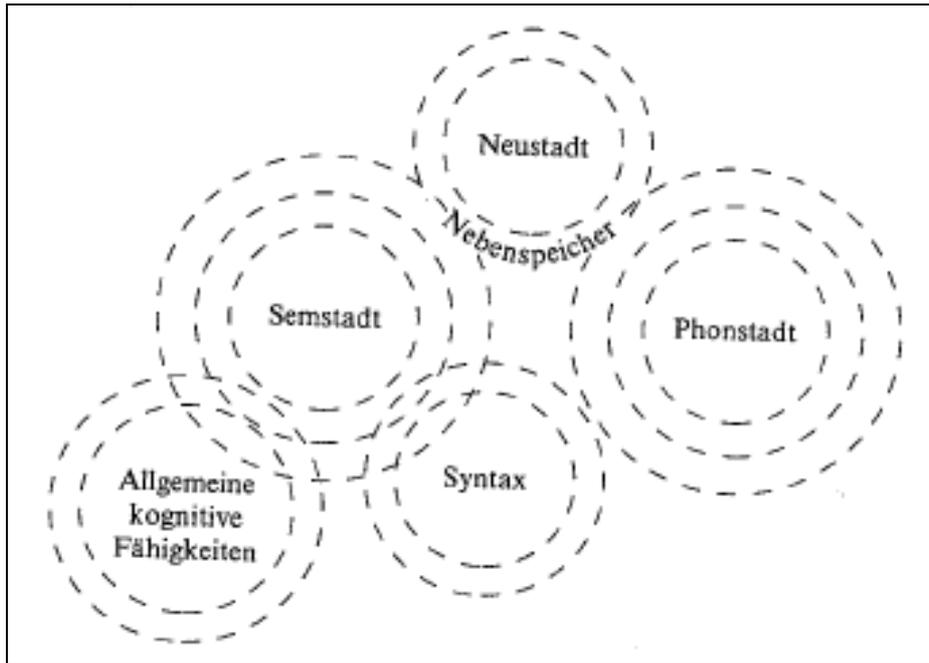


Abbildung 11: „Ein ‚Luftbild‘ der Verbindungen zu anderen Komponenten“ (Aitichson, 1997, S. 296)

4. Der kindliche Spracherwerb

Wie wir im vorigen Kapitel sehen konnten, ist die Organisation des mentalen Lexikons eine komplexe Angelegenheit. Allein das Analysieren dieser Strukturen beschäftigt die Wissenschaftler schon seit längerer Zeit. Dittmann (2002) behauptet, dass der Erwerb der Sprache die schwierigste Aufgabe in der Entwicklung der Kinder darstellt.

Laut Meibauer und Rothweiler (1999) muss ein Kind um Sprechen zu lernen „phonetisch-phonologische, semantisch-pragmatische, morphologische und syntaktische“ Aufgaben bewältigen (S. 12). Es muss sämtliche Informationen zu einem Wort aufnehmen, diese in Verbindung setzen und Assoziationen zu bereits existierenden Einträgen herstellen.

Wie Kinder den Aufbau und die Erweiterung des Wortschatzes bewerkstelligen, soll nun im folgenden Abschnitt näher erläutert werden.

4.1. Erkennen der Wortbedeutung bei Kindern

Wie im Kapitel über das mentale Lexikon schon erwähnt wurde, ist es notwendig die Bedeutungen der Worte richtig zu erfassen, um die Ausdrücke dann auch korrekt anwenden und verstehen zu können.

Aitichson (1997) beschreibt drei große Meilensteine bei der Entdeckung der Wortbedeutung im kindlichen Spracherwerb: „Sie müssen etikettieren, sortieren und ein Netzwerk bauen“ (S. 222).

4.1.1. Etikettieren

Vor dem Ende des zweiten Lebensjahres entwickelt sich laut Aitichson (1997) die Fähigkeit, bestimmte Lautkombinationen als „Bedeutung“ oder „Symbol“ eines bestimmten Objektes zu

identifizieren. Dieser Entwicklungsschritt deckt sich mit dem Einsetzen des „Wortschatzspurts“, der später noch genauer beschrieben werden soll.

Viele Eltern behaupten, dass ihre Kinder schon viel früher in der Lage sind beispielsweise bestimmte Tiere in Büchern zu benennen, wenn auch nicht mit ihrer richtigen Bezeichnung. Oftmals stellt sich aber heraus, dass diese Ausdrücke nur eine rituelle Reaktion auf möglicherweise bestimmten Seiten eines Buches darstellt. Die Einsicht, dass dieses Wort allerdings ein bestimmtes Objekt bezeichnen soll, fehlt vorerst noch.

Meibauer und Rothweiler (1999) verweisen auf eine Reihe von Prinzipien, die das Erfassen neuer Wörter erleichtern sollen:

- die Annahme, daß sich Wörter auf Klassen und nicht auf Individuen beziehen (*type assumption*, s. Clark 1993);
- die Annahme, daß alle Wörter im Lexikon einer einzigen Hierarchieebene angehören (*single level assumption*), speziell der Ebene der Basiskategorien⁸ (*basic level assumption*) (s. Clark 1993);
- die Annahme, daß sich Wörter auf ganze Objekte beziehen und nicht nur auf einen Teil oder eine Eigenschaft (*whole object assumption*, Markman 1989; Mervis 1987; *object scope principle*, Golinkoff et al. 1994);
- die Taxonomie-Annahme, derzufolge sich Wörter auf taxonomisch organisierte Kategorien beziehen (*taxonomic assumption*, Markman & Hutchinson 1984; *categorical scope principle*, Golinkoff et al. 1994);
- die Annahme, daß sich ein neues Wort auf eine bisher unbenannte Kategorie bezieht (*novel name – nameless category principle*, Golinkoff et al. 1994);

⁸ Basic-level-Ebene: „Dies ist die höchste Ebene, auf der sich ein komplexes mentales Bild formen lässt; sie enthält Einheiten wie *Hund, Hemd, Baum, Buch*.“ (Aitichson, 1997, S. 131) Dennoch ist diese Kategorie nicht eindeutig definiert.

- die Annahme, daß sich die Bedeutung von Wörtern gegenseitig ausschließt, so daß es für jedes Objekt genau einen Namen gibt (*mutual exclusivity assumption*, Markman & Wachtel 1988)
(Meibauer und Rothweiler, 1999, S.21)

4.1.2. Sortieren

Überlappend mit der Etikettierung setzt die Phase des Sortierens ein. Hier wird die Fähigkeit entwickelt einen gewissen Begriff für ein bestimmtes Objekt auf gleiche Objekte in anderen Situationen richtig anzuwenden (Aitichson, 1997).

Eine geeignete Fragestellung hierzu könnte lauten: Wie schaffen Kinder die Bezeichnung für einen bestimmten Hund, auf andere Hunde zu erweitern?

Dass dieses Herauslösen aus dem ursprünglichen Kontext oft Schwierigkeiten bereitet zeigen folgende zwei Fehler:

Unter- und Übergeneralisierung

Wenn Kinder neue Wörter lernen, kommt es häufig dazu, dass die Bedeutungen zu eng beziehungsweise manchmal aber auch zu weit gefasst werden. Nach Harris (1992, zitiert nach Meibauer und Rothweiler, 1999) ist die Untergeneralisierung die häufigere Abweichung der eigentlichen Bedeutung.

Bei der Untergeneralisierung verwenden Kinder einen Begriff oft nur für eine bestimmte Untergruppe, so kommt es beispielsweise vor, dass das Wort Hund nur für eventuell schwarze Hunde verwendet wird. Dem Kind ist somit noch nicht klar, dass das Wort Hund einen Oberbegriff darstellt, der sämtliche Hunderassen erfasst.

Die Übergeneralisierung stellt genau das gegenteilige Phänomen dar. Hier kommt es zu einer falschen Erweiterung des Begriffes. So könnten zum Beispiel alle Tiere auf vier Beinen fälschlicherweise als Hund bezeichnet werden (Meibauer und Rothweiler, 1999).

Clark (1993, zitiert nach Meibauer und Rothweiler, 1999) berichtet, dass diese zwei Arten von fehlerhafter Kategorisierung bis ins fünfte Lebensjahr beobachtet werden können.

4.1.3. Aufbau des Netzwerkes

Aitichon (1997) nimmt an, dass Wörter zunächst in einem bestimmten Kontext gelernt werden und die Verwendung dieser Worte sich auch auf diesen Kontext beschränkt. Nach und nach gliedern die Kinder das gelernte Wort ins Netzwerk ein und lösen es aus dem spezifischen Kontext. „Diese kontinuierliche Eingliederung neuer Wörter scheint ein extrem langsam voranschreitender Prozeß zu sein, der vielleicht nie abgeschlossen wird“ (Anglin, 1979, S. 99, zitiert nach Aitichon, 1997 S. 237).

Im frühen Stadium treten vor allem kollokative Verbindungen auf. Gemeint sind hier Verbindungen von Wörtern, die häufig gemeinsam auftreten. In weiterer Folge werden immer mehr Wörter ins Lexikon aufgenommen und immer mehr Verbindungen entstehen.

4.2. Die Phasen des Spracherwerbs

4.2.1. Die ersten 50 Wörter

So um den Zeitpunkt des ersten Lebensjahres beginnen Kinder ihre Lautäußerungen mit Bedeutungen zu versehen – sie erwerben ihre ersten Wörter! (Dittmann, 2002) Dieses Einsetzen der ersten Wortäußerungen unterliegt allerdings einer großen zeitlichen Variabilität. Erst wenn ein Kind im Alter von zwei Jahren die 50-Wort-Grenze noch nicht erreicht hat, besteht die Gefahr einer Sprachentwicklungsstörung, welche eine therapeutische Maßnahme erfordern würde (Grimm, 1999, zitiert nach Dittmann, 2002).

Meibauer und Rothweiler (1999) gehen von der Annahme aus, dass der kindliche Wortspeicher zunächst einmal Unterschiede in seiner Organisation und Größe aufweist. Innerhalb der nächsten paar Monate steigt der Wortschatz auf ca. 30 bis 50 Einträge an. Die Phase, in der der Umfang nur langsam wächst (ca. 2-3 Worte pro Woche), dauert meist nicht länger, als bis zur Vollendung des zweiten Lebensjahres.

Eltern versuchen ihre Kinder durch gezielte sprachliche Äußerungen beim Sprechenlernen zu unterstützen. Dittmann (2002) unterscheidet demnach folgende Stile: Ammensprache, stützende Sprache und die lehrende Sprache. Im ersten Lebensjahr der Kinder verwenden die Erwachsenen hauptsächlich die Ammensprache um mit ihren Säuglingen zu kommunizieren. Sie ist gekennzeichnet durch eine hohe Tonlage, ausgeprägte Prosodie, möglichst deutlicher Sprache mit einfachen Sätzen und Wiederholungen sowie kurze Unterbrechungen zwischen einzelnen Phrasen und der Hervorhebung besonders wichtiger Worte. Im zweiten Lebensjahr wird dieser Sprachstil von der „stützenden Sprache“ abgelöst. Bei dieser Form der Kommunikation versuchen die Erwachsenen die Aufmerksamkeit der Kinder auf bestimmte Ausschnitte zu fokussieren.

4.2.2. Der schnelle Erwerb des Wortschatzes

Ab der Hälfte des zweiten Lebensjahres setzt laut Meibauer und Rothweiler (1999) der so genannte „Wortschatzspurt“ ein. Er ist dadurch gekennzeichnet, dass in dieser Phase nun täglich mehrere Worte im mentalen Lexikon aufgenommen werden.

Fenson et al. (1994, zitiert nach Meibauer und Rothweiler, 1999) unterscheiden zwischen dem rezeptiven und dem produktiven Wortschatz, wobei der erstere deutlich größer zu sein scheint. Hat ein Kind beispielsweise die ersten 50 Wörter gelernt, so beträgt sein rezeptiver Wortschatz bereits 200 Wörter. In diesem Stadium kommt es weiters zur Ausbildung

syntaktischer⁹ Fähigkeiten und die Kinder beginnen strukturierte Kombinationen der Wörter zu verwenden. Um den Rahmen dieser Diplomarbeit allerdings nicht zu sprengen, kann auf die Entwicklung der Syntax nicht näher eingegangen werden. Hierzu verweise ich auf Clahsen (1982), der drei Theorien zum Syntaxerwerb vorstellt.

Bates et al. (1994, zitiert nach Meibauer und Rothweiler, 1999) konnten anhand einer Untersuchung mit rund 2000 englischsprachigen Kindern eine bevorzugte Aufnahme von Nomina vor Verben und anderen Prädikaten feststellen. Der Anteil der produzierten Wortklassen variiert, je nach Entwicklungsstand des Kindes deutlich. So werden im Alter von 1,5 Jahren 70,3% Substantiva und nur 12,4% Verba produziert, während sich der Anteil mit einem Alter zwischen 1,5 und 6 Jahren bei Substantiva auf 56,5% ändert und bei Verba auf 22,5% steigert (Langenmayr, 1997).

Fast mapping – Erklärungsmöglichkeit für den Wortschatzspurt

Meibauer und Rothweiler (1999) versuchen anhand des „fast mapping“ die schnelle Aufnahme von Wörtern ins mentale Lexikon zu erklären. Dieses Konzept besagt, dass Kinder schnell wichtige Informationen bezüglich der Wortform speichern und Hypothesen zu dessen Bedeutung bilden. Diese ersten Informationen zu diesem Wort sind jedoch zunächst noch sehr undifferenziert, daher wird das Wort im passiven Wortspeicher abgelegt. Sobald dieses Wort in einem anderen Kontext dargeboten wird, kann das Kind auf diesen Eintrag zugreifen und weitere Informationen dazu speichern.

Nach Crais (1992, zitiert nach Meibauer und Rothweiler 1999) sind daher mehrere Darbietungen des Wortes notwendig, um vom Kind produziert werden zu können. Folgen keine Wiederholungen, so vergisst man den Eintrag im mentalen Lexikon sehr schnell.

⁹ „Eine endliche Zahl von Einheiten einer Sprache kann durch ein System von endlich vielen Regeln, nämlich durch die Syntax der Sprache, zu potenziell unendlich vielen verschiedenen Strukturen kombiniert werden, den grammatikalisch korrekten Sätzen der Sprache.“ (Dittmann, 2002, S. 12)

Damit diese Strategie erfolgreich ist, muss daher auf die „fast mapping – Phase“ eine Phase der wiederholten Präsentation folgen.

5. Einflussfaktoren auf intellektuelle bzw. sprachliche Fähigkeiten

5.1. Alter

Die als „Defizitmodell“ bekannte Annahme, dass die Intelligenz bis zum 27. Lebensjahr ansteigt, dann langsam stagniert und kurze Zeit unverändert bleibt, bis sie dann schließlich zwischen dem 30. und 40. Lebensjahr wieder sinkt, hält in seiner allgemeinen Gültigkeit nicht Bestand. Heute weiß man, dass unterschiedliche kognitive Fähigkeiten ungleich dem Alterungsprozess unterworfen sind und eine Menge von Faktoren den Abbau und die Erhaltung der Intelligenz bedingen. Wechsler konnte feststellen, dass Untertests wie beispielsweise „Allgemeines Wissen“ und „Wortschatz“ relativ altersunabhängig waren. Hingegen Aufgaben, in denen die Geschwindigkeit ausschlaggebend war, konnten deutliche Leistungseinbußen in Abhängigkeit vom Alter beobachtet werden (Holling et al, 2004).

Kohrt & Kucharczik (1998) beschreiben ein geringeres Sprechtempo sowie häufigere Sprechpausen und eine Häufung von Versprechern bei älteren Personen. Als Gründe dafür nennen sie biologische Ursachen wie Abnahme des Hörvermögens, Verschlechterung des Kurzzeitgedächtnisses und des Konzentrationsvermögens. Anhand verschiedener Untersuchungen von Riegel (1968, zitiert nach Kohrt & Kucharczik, 1998) scheint die Identifizierung von Synonymen und Antonymen bis ins hohe Alter recht stabil zu sein. Nach Light (1993, zitiert nach Kohrt & Kucharczik, 1998) scheint die Organisation semantischer Informationen im Großen und Ganzen sehr altersstabil zu sein, allerdings können vermehrt Schwierigkeiten bei der Wortfindung festgestellt werden.

Bei der Erforschung der Intelligenzentwicklung über die Lebensspanne muss nach Funke (1998) auch mit methodischen Schwierigkeiten gerechnet werden. Vor allem bei Querschnittsstudien drängt sich die Frage auf, ob die erfassten Altersunterschiede tatsächlich

auf alterungsbedingte Ursachen zurückgeführt werden können oder ob einfach Kohorteneffekte aufgezeigt werden. Dieser Effekt kann bei verschiedenen Geburtsjahrgängen (Kohorten) aufgrund unterschiedlicher Ausbildungsbedingungen, Kriegseinflüsse, usw. auftreten. Die wesentlich aufwendigeren Längsschnittstudien scheinen auf den ersten Blick daher zuverlässigere Ergebnisse über die Intelligenzentwicklung liefern zu können. Allerdings kann es bei dieser Erhebungsmethode zu Leistungssteigerungen durch Lerneffekte bei Testwiederholungen kommen. Funke (1998) bevorzugt daher eine Kombination von Längs- und Querschnittstudien um die Intelligenzentwicklung über die Lebensspanne zu erforschen.

5.2. Geschlecht

Viele unterschiedliche Quellen stützen die Annahme, dass Frauen durchschnittlich bessere verbale Fähigkeiten besitzen als Männer. Das Konstrukt der verbalen Fähigkeiten kann laut Halpern (2000) in folgende Aspekte unterteilt werden: Word fluency, Grammatik, Rechtschreibung, Lesen, Schreiben, verbale Analogien, Wortschatz und Sprachverständnis. Die Schwierigkeit, Sprachäußerungen flüssig zu produzieren, welches als stottern bezeichnet wird, kann überdurchschnittlich oft bei Männern beobachtet werden. Gleichermäßen gilt dies für das Erscheinungsbild der Dyslexie, welche durch eine schwerwiegende Leseschwäche, bei ansonsten normalen kognitiven Fähigkeiten, beschrieben werden kann. Als Grund dafür werden biologische Faktoren angenommen, die für eine erhöhte Anfälligkeit der Männer verantwortlich sein sollen.

Nach Halpern (2000) treten erste verbale Geschlechtsunterschiede bereits in der frühen Kindheit auf. Durchschnittlich bessere Leistungen der Mädchen lassen sich vor allem beim Beginn der sprachlichen Äußerungen, bei der Länge der produzierten Äußerungen und in der Größe des Wortschatzes feststellen.

Laut Butler (1984, zitiert nach Halpern, 2000) bleibt dieser Vorsprung die ganze Volksschulzeit bestehen. Die Mehrheit der Untersuchungen belegt eine frühere Entwicklung der sprachlichen Fähigkeiten der Mädchen, allerdings scheinen diese Vorteile von geringer praktischer Relevanz zu sein.

Martin und Hoover (1987, zitiert nach Halpern, 2000) konnten anhand einer Längsschnittstudie feststellen, dass Mädchen bessere Leistungen im Rechtschreiben, Sprachgebrauch und im Leseverständnis aufweisen. Zu unterstreichen ist hier allerdings die enorme Leistungsdifferenz innerhalb der zwei Geschlechtsgruppen.

Hines (1990, zitiert nach Halpern, 2000) konnte in Ihrem Test zur Messung der „associational fluency“, bei dem zu einem vorgegeben Begriff möglichst viele Synonyme gebildet werden sollten, eine eindeutige Überlegenheit der Frauen beobachten.

Hyde und Linn (1988) postulieren hingegen, dass die Effektgrößen in unterschiedlichen verbalen Fähigkeitstests, so gering sind, dass praktisch kein Geschlechtsunterschied festgestellt werden kann. Eine einzige Ausnahme, welche eine Überlegenheit der Frauen vermuten lässt, ist die Fähigkeit der Sprachproduktion. Es scheint so, als würden Studien die vor 1973 publiziert wurden, einen größeren Unterschied zwischen Männern und Frauen hinsichtlich ihrer verbalen Fähigkeiten aufdecken.

5.3. Bildung

Zwischen den Punktwerten in Intelligenztests und bestimmten Kriterien für den Schulerfolg konnten sehr hohe Zusammenhänge gefunden werden.

Betrachtet man die höchst abgeschlossene Schulbildung in Zusammenhang mit der Intelligenz so kann ein Korrelationskoeffizient von 0,7 festgestellt werden (Wechsler, 1958, zitiert nach Amelang, 2006). Verbale Intelligenztests korrelieren höher mit schulischem Erfolg als nichtverbale Skalen, was auf eine vorwiegend sprachliche Dominanz des Unterrichts

schließen lässt. Für die Naturwissenschaften und Mathematik sind numerische Fähigkeiten von größerer Bedeutung (Amelang, 2006).

Die niedrigsten Übereinstimmungen zeigten sich jedoch beim Vergleich von Intelligenz mit dem Studienerfolg. Als Gründe dafür können die geringe Leistungsvariabilität unter den Studenten sowie eine außerordentlich niedrige psychometrische Qualität der Noten genannt werden (Amelang, 2006).

Zu diesem Ergebnis kam auch Jensen (1998, zitiert nach Holling et al, 2004), der den Zusammenhang zwischen Intelligenz und Schulleistung bei unterschiedlichen Schultypen untersuchte. Er konnte feststellen, dass der Zusammenhang bei steigendem Ausbildungsniveau sinkt. Anhand von Studien im amerikanischen Raum ergab sich in der Grundschule eine Korrelation zwischen $r=0,60$ und $r=0,70$; für die Highschool zwischen $r=0,50$ und $r=0,60$; für das College zwischen $r=0,40$ und $0,50$ und für Graduate School zwischen $r=0,30$ und $r=0,40$.

III. EMPIRISCHER TEIL

6. Testmaterial und Durchführung der Testung

Ziel dieser Arbeit ist die Evaluation zweier, von Dr. Arendasy konstruierten Sprachtests zur Messung der Wortflüssigkeit (WF engl) und des Wortschatzes (Synonyme Fragebogen engl-Version 1) anhand englischsprachiger Testpersonen (native speakers). Die Testung wurde zwischen 24. April und 9. Juli 2008 als Online-Studie durchgeführt und konnte von den Testpersonen mittels Zugangscodes unter folgendem Link

<http://131.130.64.42/diffpsylabor/start.aspx> (Stand 10.9.2008) aufgerufen werden.

Dieses Testweb wurde 2002 von Dr. Arendasy entwickelt und ermöglicht den Zugriff auf sämtliche Tests über das Internet. Daher konnten auch Personen im Ausland an dieser Studie problemlos teilnehmen, wodurch die Anforderungen an eine Testgruppe mit Englisch als Muttersprache leichter realisiert werden konnte. Die Probanden wurden einerseits persönlich und andererseits per E-mail kontaktiert. Nachdem die Instruktionen im Testweb selbst leider nur auf Deutsch bereitgestellt werden konnten, wurde zusätzlich eine englische Anleitung verteilt, um sicherzustellen, dass die Aufgaben auch verstanden wurden.

Im Zuge der Testung wurden zunächst persönliche Daten erhoben (Alter, Geschlecht und höchst abgeschlossene Ausbildung). Um die Anonymität der teilnehmenden Personen zu gewährleisten, wurden Sie gebeten einen Probandencode festzulegen. Dieser sollte in beiden Tests gleich gewählt werden, um die Leistungen in beiden Tests später vergleichen zu können. In beiden Verfahren wurden die teilnehmenden Personen instruiert die Aufgaben durchgehend, konzentriert, zügig und ohne jegliche Hilfsmittel zu bearbeiten.

Synonyme Fragebogen engl (Version 1)

Bei dem hier entwickelten Test handelt es sich um ein Verfahren zur Messung des Wortschatzes. Die Aufgabe bestand darin, zu einem vorgegebenen Begriff ein möglichst bedeutungsähnliches Wort aus vier Antwortalternativen zu finden.

Beispiel Item 1:

The term „clock“ is most similar to the term

- regulator
- timekeeper
- measurement device
- hygrometer

Dieses Verfahren kann als reiner Power-Test angesehen werden, da keine enge Zeitbeschränkung vorgegeben ist.

Eine Studie zur deutschen Version (Synonyme-Fragebogen – Version 1) findet sich bei Huber (2008).

Word Fluency Test „WF engl“

Dieser Test, bestehend aus 20 Items, erfordert von den Testpersonen die Fähigkeit, aus einer durcheinander geratenen Buchstabenfolge ein Wort zu erkennen. Durch die richtige Reihung der Buchstaben ergibt sich ein sinnvolles, englisches Wort, welches aus der Alltagssprache bekannt sein sollte. Die Aufgabe bestand nun darin, dieses Wort zu erkennen und in das freie Feld einzutragen. Kann jedoch kein sinnvolles Wort abgeleitet werden, so erhält die Testperson die Instruktion dieses Feld frei zu lassen und den Button „Weiter“ zu drücken.

Um sicherzustellen, dass die Testpersonen die Instruktion verstanden haben, werden Ihnen vor der eigentlichen Testbearbeitung zwei Übungsaufgaben vorgegeben.

Beispielaufgabe 2: PSYHCLOOGIE
Lösung: PSYCHOLOGIE

Es wurde darauf hingewiesen, dass Sie sich genug Zeit für die Bearbeitung des Tests nehmen und Unterbrechungen vermeiden sollten. Da auch bei diesem Verfahren keine enge Zeitbeschränkung vorgegeben ist, kann er ebenfalls als Power-Test eingestuft werden. Eine Studie zur deutschen Version des Wortflüssigkeitstests kann bei Gallistl (2006) nachgelesen werden.

Nachfolgende Berechnungen wurden mittels SPSS 10.0 durchgeführt. Zur Überprüfung der Rasch-Homogenität wurde das Programm LPCMWIN verwendet.

7. Synonyme Fragebogen engl (Version 1)

7.1. Beschreibung der Stichprobe

7.1.1. Auswahl der gültigen Testpersonen

Die Gesamtstichprobe in diesem Test umfasst zunächst 83 Personen, wobei 10 Testpersonen in weiterer Folge ausgeschlossen werden mussten. Das Ausschlusskriterium war dabei eine unglaubliche Bearbeitungszeit. Alle Personen, die eine Testbearbeitungsdauer von unter 60 Sekunden und über 1800 (entspricht 30 Minuten) aufwiesen, wurden von den weiteren Analysen ausgeschlossen, da angenommen werden muss, dass eine ernsthafte Bearbeitung der Items in dieser Zeit nicht möglich ist. Alle Berechnungen zum Test „Synonyme Fragebogen engl (Version 1) beziehen sich somit auf die bereinigte Stichprobe, die nunmehr 73 Personen beinhaltet. Es wurden rund 12% von der Gesamtstichprobe entfernt.

7.1.2. Bearbeitungsdauer

Anhand der gültigen Testpersonen kann eine minimale Bearbeitungszeit von 60 Sekunden und eine maximale Testdauer von 1075 Sekunden (entspricht 17 Minuten und 55 Sekunden) beobachtet werden. Der Median beträgt 233 Sekunden (entspricht 3 Minuten und 53 Sekunden) der Mittelwert (MW) ist 288,18 (entspricht 4 Minuten und 48 Sekunden) bei einer Standardabweichung (SD) von 190,82 (entspricht 3 Minuten und 11 Sekunden).

In der hier dargestellten Grafik (vgl. Abbildung 12) wurden die Bearbeitungszeiten in Gruppen (alle 60 Sekunden) zusammengefasst d.h. Gruppe 1 = 1 – 60 sec. (bis 1 Minute); 2 = 61 – 120 sec. (bis 2 Minuten) usw.

Tabelle 1: Deskriptivstatistische Kennzahlen zur Testbearbeitungsdauer im Synonyme-Test (N = 73)

| | MW | SD | Median | Minimum | Maximum |
|-------------------------|--------|--------|--------|---------|---------|
| Testdauer (sec.) | 288,18 | 190,82 | 233 | 60 | 1075 |

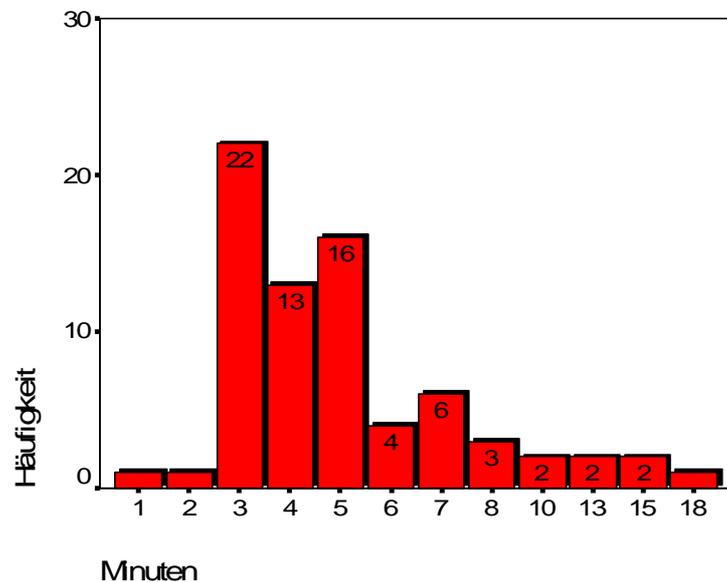


Abbildung 12: Bearbeitungsdauer im Synonyme-Test in Minuten (N = 73)

7.1.3. Alter

Die anfallende Stichprobe setzt sich zusammen aus drei Testpersonen zwischen 11 und 20 Jahren (4,1%), 33 Personen zwischen 21 und 30 Jahren (45,2%), 14 getesteten Personen zwischen 31 und 40 Jahren (19,2%), sieben Personen zwischen 41 und 50 Jahren (9,6%), 13 Testpersonen zwischen 51 und 60 Jahren (17,8%) und drei getesteten Personen zwischen 61 und 70 Jahren (4,1%). Die jüngste Person ist 19 Jahre alt, während der älteste Teilnehmer 69 Jahre alt ist. Das Durchschnittsalter (MW) kann mit rund 35,36 Jahren (SD = 13,79) angegeben werden, wobei der Median hier bei 31 Jahren liegt.

Tabelle 2: Deskriptivstatistische Kennzahlen zum Alter der Testpersonen im Synonyme-Test (N = 73)

| | MW | SD | Median | Minimum | Maximum |
|--------------------------|-------|-------|--------|---------|---------|
| Alter (in Jahren) | 35,36 | 13,79 | 31 | 19 | 69 |

Tabelle 3: Verteilung und Anteilswerte des Lebensalters im Synonyme-Test (N = 73)

| Alter | Häufigkeit | Prozent |
|---------------|-------------------|----------------|
| 11-20 | 3 | 4,1 |
| 21-30 | 33 | 45,2 |
| 31-40 | 14 | 19,2 |
| 41-50 | 7 | 9,6 |
| 51-60 | 13 | 17,8 |
| 61-70 | 3 | 4,1 |
| Gesamt | 73 | 100 |

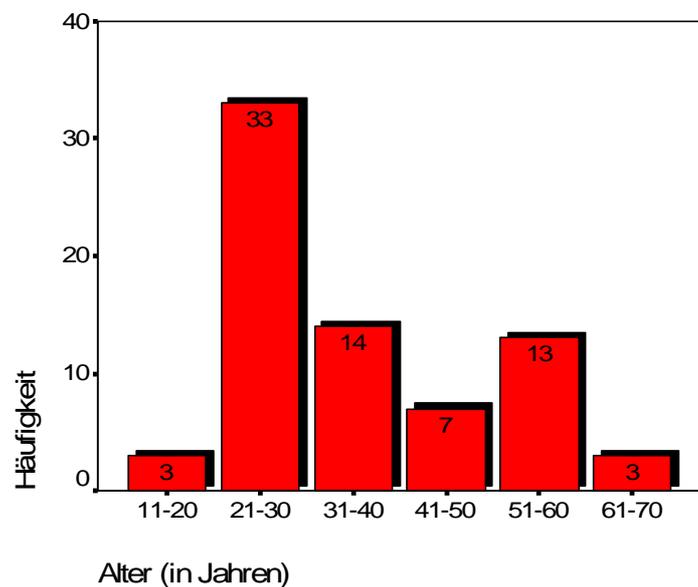


Abbildung 13: Verteilung des Lebensalters im Synonyme-Test (N= 73)

7.1.4. Geschlecht

Die vorliegende Stichprobe umfasst 31 männliche und 42 weibliche Teilnehmer, was einem Anteil von 42,5 % Männer und 57,5 % Frauen entspricht.

Tabelle 4: Verteilung und Anteilswerte des Geschlechts im Synonyme-Test (N = 73)

| Geschlecht | Häufigkeit | Prozent |
|---------------|------------|------------|
| männlich (m) | 31 | 42,5 |
| weiblich (w) | 42 | 57,5 |
| Gesamt | 73 | 100 |

7.1.5. Bildung

Aufgrund der Schwierigkeit von Österreich aus möglichst viele englischsprachige Testpersonen zu finden, wurde eine anfallende Stichprobe gewählt. Die Testpersonen wurden gebeten, die höchst abgeschlossene Ausbildung anzugeben, wobei folgenden Auswahlmöglichkeiten zur Verfügung standen:

Tabelle 5: Verteilung und Anteilswerte der höchsten abgeschlossenen Ausbildung im Synonyme-Test (N = 73)

| Kategorie | Höchste abgeschlossene Ausbildung | Häufigkeit | Prozent |
|-----------|---|------------|---------|
| 0 | Volksschule = Primary education (elementary school) Sonderschule = special school | 2 | 2,7 |
| 1 | Hauptschule / AHS Unterstufe = secondary education first level (middle school / Junior High School) | 1 | 1,4 |
| 2 | Fachschule = vocational school; Lehre = apprenticeship | 0 | 0 |
| 3 | Matura = Secondary education second level (Senior High School) | 9 | 12,3 |
| 4 | Fachhochschule = Post secondary education (Bachelor's degree; university applied science); Akademie = academy | 9 | 12,3 |
| 5 | Universität = First stage of tertiary education – university (Master's Degree Study) | 48 | 65,8 |
| 6 | Doktoratsstudium = Secondary stage of tertiary education – university (Doctor's Degree Study) | 4 | 5,5 |

Betrachtet man die Verteilung, so kann festgestellt werden, dass zwei Testpersonen eine Volksschule oder eine Sonderschule besuchten (2,7%), eine Person die Hauptschule oder AHS Unterstufe absolvierte (1,4%), keiner eine Fachschule oder Lehre, neun Personen die Matura (12,3%), neun Testpersonen eine Fachhochschule oder Akademie (12,3%), 48 Personen haben eine Universität mit Magister abgeschlossen (65,8%) und vier der getesteten Personen haben das Doktoratsstudium beendet (5,5%). Aufgrund der überwiegenden Anzahl von Testpersonen mit höherer Ausbildung ist diese Stichprobe bezüglich des Ausbildungsgrades als überdurchschnittlich gebildet zu charakterisieren.

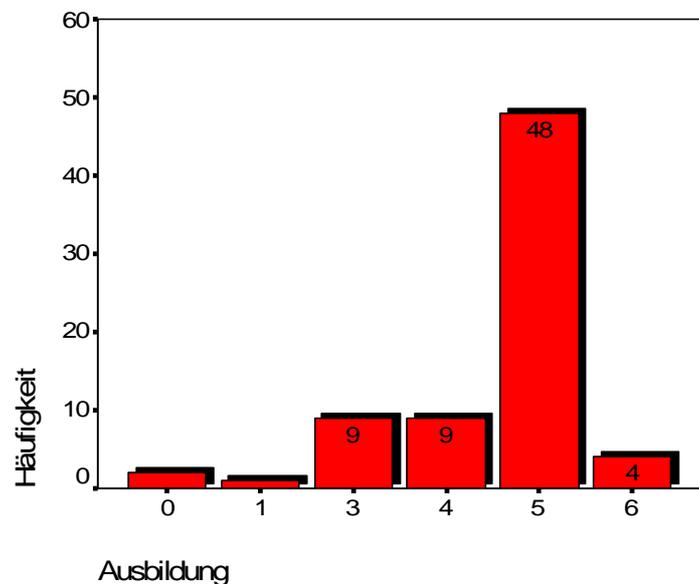


Abbildung 14: Verteilung der höchsten abgeschlossenen Ausbildung im Synonyme-Test (N = 73)

7.2. Reliabilität und Trennschärfen (N = 73)

Für den ursprünglichen Gesamttest ($k = 20$) würde sich ein Cronbach-Alpha von 0,30 ergeben. Durch Entfernung der negativ-trennscharfen Items (1, 12, 14, 17 und 19) sowie des zu leichten Items 16, verbleiben schließlich 14 Items, die dann ein Cronbach-Alpha von 0,42 und Trennschärfen zwischen $r = 0,00$ und $r = 0,32$ aufweisen.

Kubinger (2006) betont die Forderungen der Klassischen Testtheorie, dass sich die Itemschwierigkeiten gleichmäßig innerhalb des Intervalls von 0,05 und 0,95 verteilen sollten. Die mit (*) gekennzeichneten Items (vgl. Tabelle 6) werden aufgrund negativer Trennschärfen oder zu hoher Lösungswahrscheinlichkeit ($p_i \geq 0,95$) eliminiert.

Tabelle 6: Itemkennwerte im Synonyme-Test (k= 20)

| Item | Lösungswahrscheinlichkeit (p_i) | Standardabweichung (σ_i) | Korrigierte Itemtrennschärfe $\rho(x_i, X^*)$ | Alpha (if item deleted) |
|------|-------------------------------------|-----------------------------------|---|-------------------------|
| 1* | 0,75 | 0,43 | -0,48* | 0,33 |
| 2 | 0,92 | 0,28 | 0,74 | 0,29 |
| 3 | 0,08 | 0,28 | 0,28 | 0,25 |
| 4 | 0,81 | 0,40 | 0,41 | 0,18 |
| 5 | 0,44 | 0,50 | 0,03 | 0,30 |
| 6 | 0,79 | 0,41 | 0,06 | 0,29 |
| 7 | 0,95 | 0,23 | 0,20 | 0,27 |
| 8 | 0,37 | 0,49 | 0,13 | 0,27 |
| 9 | 0,75 | 0,43 | 0,11 | 0,28 |
| 10 | 0,90 | 0,30 | 0,05 | 0,29 |
| 11 | 0,74 | 0,44 | 0,22 | 0,24 |
| 12* | 0,84 | 0,37 | -0,04* | 0,32 |
| 13 | 0,34 | 0,48 | 0,05 | 0,30 |
| 14* | 0,81 | 0,40 | -0,09* | 0,33 |
| 15 | 0,30 | 0,46 | 0,18 | 0,25 |
| 16* | 0,99* | 0,12 | 0,06 | 0,29 |
| 17* | 0,92 | 0,28 | -0,02* | 0,31 |
| 18 | 0,68 | 0,47 | 0,05 | 0,30 |
| 19* | 0,15 | 0,36 | -0,12* | 0,34 |
| 20 | 0,55 | 0,50 | 0,09 | 0,28 |

7.3. Verteilung des Rohscores

Aufgrund der vorangegangenen Berechnungen (siehe Itemtrennschärfe und Itemschwierigkeiten) mussten 6 Items ausgeschlossen werden. Die hier berechnete Verteilung

des Rohscores bezieht sich somit auf die verbleibenden 14 Items. Es kann ein minimaler Rohscore von 4 und ein maximaler Wert von 13 richtig gelösten Items abgelesen werden. Der Mittelwert (MW) liegt bei 8,63 richtigen Lösungen bei einer Standardabweichung (SD) von 1,99. Der Median beträgt in dieser Stichprobe 8.

Tabelle 7: Deskriptivstatistischen Kennzahlen zum Rohwert im Synonyme-Test (k = 14; N = 73)

| | Median | MW | SD | Minimum | Maximum |
|-----------------|---------------|-----------|-----------|----------------|----------------|
| Rohscore | 8 | 8,63 | 1,99 | 4 | 13 |

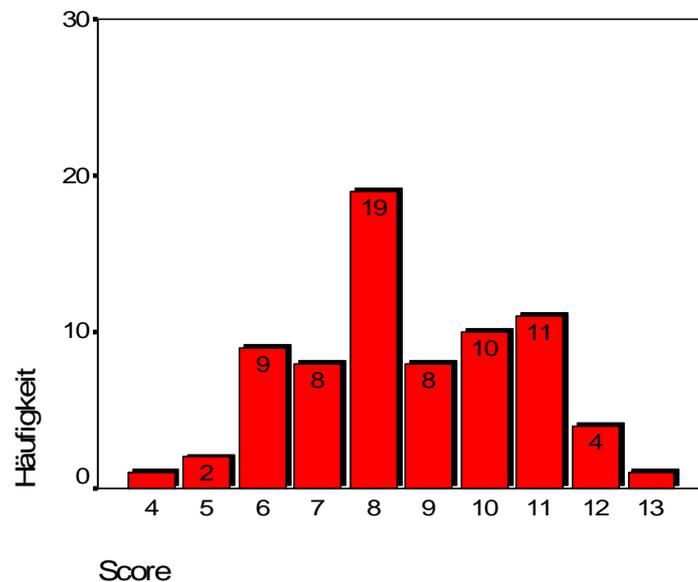


Abbildung 15: Verteilung des Rohscores im Synonyme-Test (k = 14, N = 73)

Die Verteilung des Rohscores im Synonyme-Test, untersucht mittels Kolmogorov-Smirnov-Test ergibt $z = 1,35$ und $p = 0,051$. Die Normalverteilung der Daten kann somit angenommen werden.

7.4. Überprüfung auf Rasch-Homogenität

Die Rasch-Homogenität ist von großer Bedeutung, weil diese Analyse hypothesentestend untersucht, ob die Items dieselbe Fähigkeit, also eindimensional, messen. Durch einen einzigen Parameter kann angegeben werden, welchen Grad der Schwierigkeit das Item aufweist. Der Vorteil des Rasch-Modells ist, dass die Fähigkeit der Person als auch die Schwierigkeit des Items eindimensional erfasst werden kann (Kubinger, 2006).

Von den ursprünglich $k = 20$ Items wurden 6 Items (Item 1, 12, 14, 16, 17, 19) aufgrund ihrer negativen Trennschärfe bzw. zu hohen Lösungswahrscheinlichkeit bereits ausgeschlossen.

Mit den 14 verbliebenen Items wird nun eine Überprüfung auf Rasch-Homogenität durchgeführt. Der Stichprobenumfang setzt sich aus jenen 73 Testpersonen zusammen, die den Test instruktionsgemäß bearbeitet haben.

Tabelle 8: Schwierigkeitsparameter und Standardfehler der einzelnen Items im Synonyme-Test (Gesamtstichprobe)

| Item | Itemparameter | Standardfehler |
|------|-----------------|----------------|
| 1* | Nicht verwendet | |
| 2 | -1,88 | 0,40 |
| 3 | +3,21 | 0,40 |
| 4 | -0,88 | 0,26 |
| 5 | +0,93 | 0,15 |
| 6 | -0,79 | 0,22 |
| 7 | -2,32 | 0,48 |
| 8 | +1,24 | 0,17 |
| 9 | -0,54 | 0,20 |
| 10 | -1,71 | 0,37 |
| 11 | -0,47 | 0,18 |
| 12* | Nicht verwendet | |
| 13 | +1,37 | 0,17 |
| 14* | Nicht verwendet | |
| 15 | +1,57 | 0,18 |
| 16* | Nicht verwendet | |
| 17* | Nicht verwendet | |
| 18 | -0,18 | 0,23 |

| | | |
|----|-----------------|------|
| 19 | Nicht verwendet | |
| 20 | +0,45 | 0,23 |

Anmerkung: Die Itemparameter in der Gesamtstichprobe wurden hier mit (-1) multipliziert.

Positive Werte sind hier als schwierigere Items zu verstehen, negative Werte daher als leichtere Items. Es gilt, dass die Summe der Schwierigkeitsparameter über die 14 Items Null ergibt (Nullsummennormierung, $\sum \sigma_i = 0$).

Eine weitere besondere Eigenschaft des Rasch-Modells ist seine Stichprobenunabhängigkeit.

Diese gestattet es, die Stichprobe anhand bestimmter Merkmale in Teilstichproben zu unterteilen.

7.4.1. Teilungskriterium Geschlecht

Die Prüfung auf Rasch-Konformität der Items im Synonyme-Test wird zunächst anhand des externen Teilungskriteriums Geschlecht (weiblich, männlich) durchgeführt.

Von den 73 gültigen Testpersonen sind 42 weiblich und 31 männlich.

Als globale Testgröße im Likelihood-Quotiententest wird Chi-Quadrat (χ^2) nach Andersen berechnet, die Schwierigkeitsparameter aus den Teilstichproben werden zur Prüfung der Modellgeltung für alle Items ($df = k-1$) gemeinsam berücksichtigt (die Chi-Quadrat-Approximation wird nach Wilson-Hilferty durchgeführt). Es lassen sich die Likelihoods von zwei unterschiedlichen Modellen miteinander vergleichen (Rost, 2004). Ist der beobachtete Wert größer als der kritische χ^2 – Wert, dann ist eine Rasch-Homogenität nicht anzunehmen; die Daten sind dann durch die getrennt geschätzten Itemparameter besser beschreibbar.

Tabelle 9: Likelihood-Quotiententest mittels χ^2 nach Andersen im Synonyme-Test (Teilungskriterium Geschlecht)

| Teilungskriterium Geschlecht | |
|------------------------------|-------|
| Andersen - χ^2 | 9,93 |
| Freiheitsgrade (df) | 13 |
| χ^2 bei alpha 5% | 22,35 |
| χ^2 bei alpha 1% | 27,72 |

Die Prüfung fällt sowohl bei alpha = 5% als auch 1% nicht signifikant aus, die Modellgeltung nach dem Teilungskriterium Geschlecht kann als gegeben angenommen werden.

Die einzelnen Items werden im Detail zur Beurteilung der Signifikanz der

Parameterschätzwertedifferenz mittels Wald-Statistik überprüft (Kritische z-Werte für alpha 5% = 1,96, für alpha 1% = 2,58).

Tabelle 10: Itemparameter der Teilstichproben („weiblich“ und „männlich“) mit Signifikanzbeurteilung nach Wald im Synonyme-Test

| Item | Itemparameter „weiblich“ | Itemparameter „männlich“ | z - Wert | Signifikanz |
|------|--------------------------|--------------------------|----------|-------------|
| 2 | -2,05 | -1,62 | +0,52 | n.s. |
| 3 | +3,25 | +3,25 | +0,01 | n.s. |
| 4 | -1,08 | -0,56 | +1,00 | n.s. |
| 5 | +1,11 | +0,74 | -0,92 | n.s. |
| 6 | -1,08 | -0,37 | +1,67 | n.s. |
| 7 | -2,05 | -2,82 | -0,91 | n.s. |
| 8 | +1,00 | +1,66 | +1,86 | n.s. |
| 9 | -0,37 | -0,77 | -0,78 | n.s. |
| 10 | -1,27 | -2,82 | -2,09 | sign. (5%). |
| 11 | -0,50 | -0,37 | +0,29 | n.s. |
| 13 | +1,44 | +1,34 | -0,27 | n.s. |
| 15 | +1,69 | +1,49 | -0,52 | n.s. |
| 18 | -0,26 | -0,03 | +0,51 | n.s. |
| 20 | +0,18 | +0,88 | +2,02 | sign. (5%) |

Die Überprüfung der einzelnen Items ergibt, dass nur die Items 10 und 20 in den Teilstichproben als unterschiedlich schwierig zu bezeichnen sind.

7.4.2. Teilungskriterium Alter

Als weiteres externes Teilungskriterium wird nun das Alter herangezogen. Der Cutting-Score der Stichprobe im Synonyme-Test ist der Mittelwert des Lebensalters, Testpersonen bis 34 Jahre werden somit als „jung“, jene ab 35 Jahren als Teilstichprobe „alt“ bezeichnet. Daher können 41 Personen (56,2%) als „jung“ und 32 Personen (43,8%) als „alt“ klassifiziert werden.

Tabelle 11: Likelihood-Quotiententest mittels χ^2 nach Andersen im Synonyme-Test (Teilungskriterium Alter)

| Teilungskriterium Alter | |
|--------------------------------|-------|
| Andersen - χ^2 | 10,33 |
| Freiheitsgrade (df) | 13 |
| χ^2 bei alpha 5% | 22,35 |
| χ^2 bei alpha 1% | 27,72 |

Die Prüfung fällt sowohl bei alpha = 5% als auch 1% nicht signifikant aus, die Modellgeltung nach dem Teilungskriterium Alter kann jeweils als gegeben angenommen werden.

Die einzelnen Items werden im Detail zur Beurteilung der Signifikanz der Parameterschätzwertedifferenz mittels Wald-Statistik überprüft (Kritische z-Werte für alpha 5% = 1,96, für alpha 1% = 2,58).

Tabelle 12: Itemparameter der Teilstichproben („jung“ und „alt“) mit Signifikanzbeurteilung nach Wald im Synonyme-Test

| Item | Itemparameter „jung“ | Itemparameter „alt“ | z - Wert | Signifikanz |
|------|----------------------|---------------------|----------|----------------|
| 2 | -2,19 | -1,33 | 1,07 | n.s. |
| 3 | 2,76 | 3,84 | 1,27 | n.s. |
| 4 | -0,77 | -1,00 | -0,40 | n.s. |
| 5 | 0,90 | 1,03 | 0,32 | n.s. |
| 6 | -0,39 | -1,77 | -2,60 | sign. (5%; 1%) |
| 7 | -2,63 | -1,77 | 1,08 | n.s. |
| 8 | 1,12 | 1,44 | 0,85 | n.s. |
| 9 | -0,63 | -0,32 | 0,78 | n.s. |
| 10 | -1,41 | -2,50 | -1,08 | n.s. |
| 11 | -0,51 | -0,32 | 0,48 | n.s. |
| 13 | 1,49 | 1,30 | -0,49 | n.s. |
| 15 | 1,63 | 1,58 | -0,12 | n.s. |
| 18 | 0,05 | -0,51 | -1,07 | n.s. |
| 20 | 0,57 | 0,34 | -0,56 | n.s. |

Die Überprüfung der einzelnen Items ergibt, dass nur Item 6 in diesen Teilstichproben als unterschiedlich schwierig zu bezeichnen ist.

7.4.3. Teilungskriterium Rohscore

Als internes Teilungskriterium, wird die Leistung im Test (niedriger vs. hoher Rohscore) selbst herangezogen. Das interne Kriterium erweist sich zumeist als besonders sensitiv, im Gegensatz zu externen Teilungskriterien.

Die Stichprobe wird aufgrund des Mittelwertes der Testleistung der Personen (MW = 8,63; Md = 8) geteilt: Die Teilstichproben setzen sich aus Personen mit „wenigen“ richtigen Antworten (Score \leq 8) und jenen mit mehr korrekten Antworten (Score $>$ 8) zusammen. Insofern haben 39 Personen (53,4%) wenige, 34 Personen (46,6%) mehr Items gelöst.

Tabelle 13: Likelihood-Quotiententest mittels χ^2 nach Andersen im Synonyme-Test (Teilungskriterium Rohscore)

| Internes Teilungskriterium Rohscore (MW) | |
|---|-------|
| Andersen - χ^2 | 16,50 |
| Freiheitsgrade (df) | 13 |
| χ^2 bei alpha 5% | 22,35 |
| χ^2 bei alpha 1% | 27,72 |

Die Prüfung fällt sowohl bei alpha = 5% als auch 1% nicht signifikant aus, die Modellgeltung nach dem Teilungskriterium Rohscore kann zunächst als gegeben angenommen werden. Die Überprüfung der Parameterschätzwertedifferenz auf Signifikanz ergibt allerdings pro Item, dass diese nicht schätzbar sind. Diese Beobachtung lässt darauf schließen, dass einerseits die Differenzen der Itemparameter beträchtlich sind und andererseits, dass der Stichprobenumfang in diesem Fall keine genaue Schätzung gestattet.

Auch weist der grafische Modelltest darauf hin, dass die Items weit außerhalb der 45°-Geraden zu liegen kommen. Dieser Umstand kann als Diskrepanz zum nicht-signifikanten Ergebnis im Andersen- χ^2 -Test gewertet werden, sodass das interne Teilungskriterium im Synonyme-Test auf eine Verletzung der Rasch-Homogenität hindeutet.

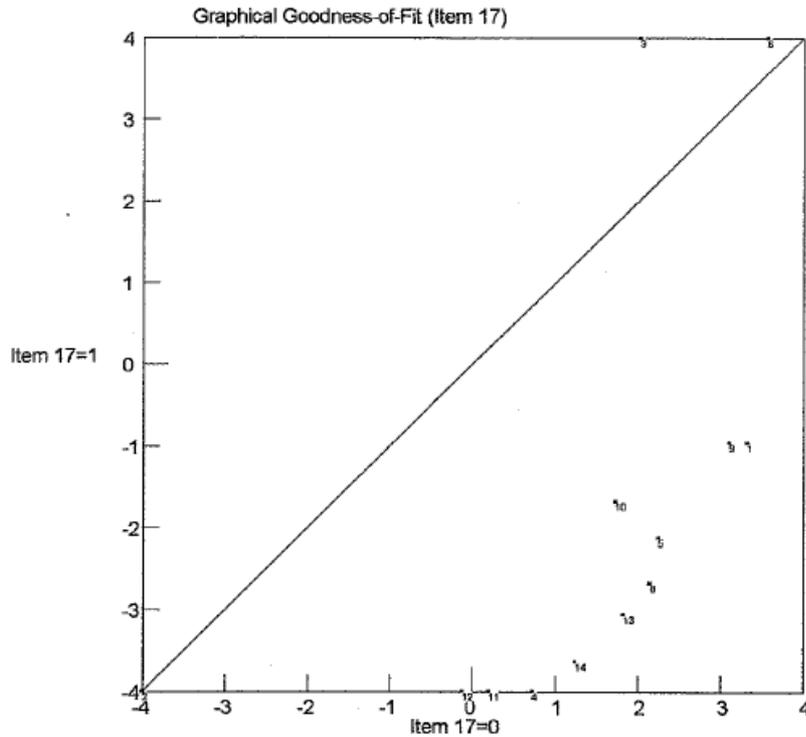


Abbildung 16: Grafische Modellprüfung im Synonyme-Test internes Teilungskriterium ($k = 14$; $N = 73$)

7.4.4. Personenparameter im Synonyme-Test

Die erschöpfende Statistik im Sinne des Rasch-Modells wird als Berechtigung verstanden, dass der Summenscore pro Testperson eine suffiziente Statistik für den Leistungsgrad der Person ist. Es ist daher nicht mehr wichtig, welche Items gelöst wurden, sondern nur mehr wie viele (Rost, 2004).

Anhand der Datenanalyse geht im Sinne der erschöpfenden Statistiken hervor, dass Personen mit einer bestimmten Anzahl gelöster Items im Synonyme-Test einen eindeutigen Fähigkeitsparameter (Θ) erhalten.

Tabelle 14: Rohwerte und die entsprechenden Fähigkeitsparameter (Θ) im Synonyme-Test

| Rohwert | Theta (Θ) |
|---------|--------------------|
| 1 | -3,32 |
| 2 | -2,43 |
| 3 | -1,83 |
| 4 | -1,33 |
| 5 | -0,88 |
| 6 | -0,46 |
| 7 | -0,05 |
| 8 | +0,37 |
| 9 | +0,80 |
| 10 | +1,27 |
| 11 | +1,81 |
| 12 | +2,47 |
| 13 | +3,46 |

Im vorliegenden Synonyme-Test haben die Personen zwischen 4 und 13 Items gelöst. Der Fall, dass Personen keine oder alle Items gelöst haben, kommt in dieser Stichprobe nicht vor; d.h. die zugehörigen Parameter $\Theta_{r=0} = -\infty$ und $\Theta_{r=k} = +\infty$ sind nicht erforderlich.

8. Word Fluency Test „WF engl“

8.1. Beschreibung der Stichprobe

8.1.1. Auswahl der gültigen Testpersonen

Insgesamt bearbeiteten zunächst 42 Personen den Wortflüssigkeitstest (WF engl). Aufgrund der zu geringen, weil unglaublich kurzen Testdauer, mussten auch hier fünf Testpersonen von den weiteren Berechnungen ausgeschlossen werden. Das Ausschlusskriterium war eine Bearbeitungszeit unter 120 Sekunden. Nachdem auch hier rund 11,9% der Gesamtstichprobe entfernt wurden, verbleibt eine bereinigte Stichprobe von 37 Testpersonen.

8.1.2. Bearbeitungsdauer

Anhand der verbleibenden 37 Personen kann eine minimale Testbearbeitungszeit von 155 Sekunden (entspricht 2 Minuten und 35 Sekunden), sowie eine maximale Dauer von 3189 Sekunden (entspricht 53 Minuten und 9 Sekunden) angegeben werden. Die durchschnittliche Testdauer (MW) beträgt 876,11 Sekunden (entspricht 14 Minuten und 36 Sekunden) bei einer Standardabweichung (SD) von 721,78 Sekunden (entspricht 12 Minuten und 2 Sekunden).

Der Median beträgt 693 Sekunden (entspricht 11 Minuten und 33 Sekunden)

In der hier dargestellten Grafik (vgl. Abbildung 17) wurden die Bearbeitungszeiten in Gruppen (alle 60 Sekunden) zusammengefasst d.h. Gruppe 1 = 1 – 60 sec. (bis 1 Minute); 2 = 61 – 120 sec. (bis 2 Minuten) usw.

Tabelle 15: Deskriptivstatistische Kennzahlen zur Testbearbeitungsdauer im Wortflüssigkeitstest (N = 37)

| | MW | SD | Median | Minimum | Maximum |
|-------------------------|--------|--------|--------|---------|---------|
| Testdauer (sec.) | 876,11 | 721,78 | 693 | 155 | 3189 |

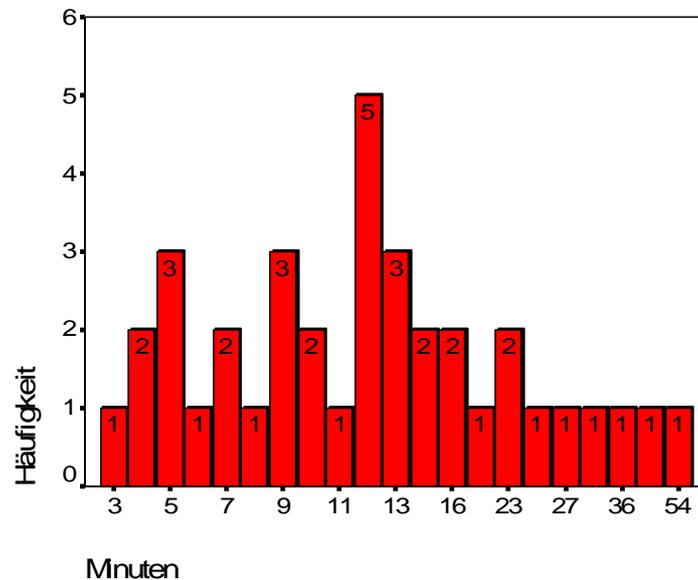


Abbildung 17: Bearbeitungsdauer im Wortflüssigkeitstest in Minuten (N = 37)

8.1.3. Alter

Die hier vorliegende Stichprobe setzt sich zusammen aus einer Person zwischen 10 und 20 Jahren (2,7%), 19 Testpersonen zwischen 21 und 30 Jahren (51,4%), sieben Teilnehmern zwischen 31 und 40 Jahren (18,9%), fünf Personen zwischen 41 und 50 Jahren (13,5 %) und ebenfalls fünf getesteten Personen zwischen 51 und 60 Jahren (13,5%). Die jüngste Testperson in dieser Stichprobe ist 19 und die älteste 59 Jahre alt. Das Durchschnittsalter (MW) kann mit 33,11 Jahren angegeben werden, bei einer Standardabweichung von 11,76. Der Median liegt bei 27 Jahren liegt.

Tabelle 16: Deskriptivstatistische Kennzahlen zum Alter der Testpersonen im Wortflüssigkeitstest (N = 37)

| | MW | SD | Median | Minimum | Maximum |
|--------------------------|-------|-------|--------|---------|---------|
| Alter (in Jahren) | 33,11 | 11,76 | 27 | 19 | 59 |

Tabelle 17: Verteilung und Anteilswerte des Lebensalters im Wortflüssigkeitstest (N = 37)

| Alter | Häufigkeit | Prozent |
|---------------|------------|------------|
| 11-20 | 1 | 2,7 |
| 21-30 | 19 | 51,4 |
| 31-40 | 7 | 18,9 |
| 41-50 | 5 | 13,5 |
| 51-60 | 5 | 13,5 |
| 61-70 | 0 | 0 |
| Gesamt | 37 | 100 |

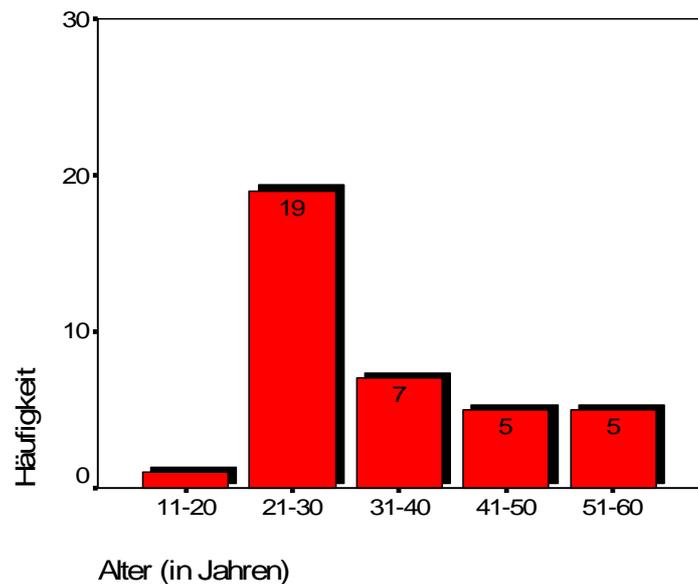


Abbildung 18: Verteilung des Lebensalters im Wortflüssigkeitstest (N= 37)

8.1.4. Geschlecht

Die Geschlechterverteilung in der erhobenen Stichprobe kann als sehr ausgewogen beschrieben werden, 18 weibliche Testpersonen (48,6%) und 19 Männer (51,4%) nahmen an der Studie teil.

Tabelle 18: Verteilung und Anteilswerte des Geschlechts im Wortflüssigkeitstest (N = 37)

| Geschlecht | Häufigkeit | Prozent |
|-------------------|-------------------|----------------|
| männlich | 19 | 51,4 |
| weiblich | 18 | 48,6 |
| Gesamt | 37 | 100 |

8.1.5. Bildung

Deskriptivstatistische Analysen ergaben, dass drei der getesteten Personen eine Volksschule oder Sonderschule absolvierten (8,1%), keiner eine Hauptschule oder AHS Unterstufe als höchst abgeschlossene Ausbildung angab, keiner eine Fachschule oder Lehre, zwei Personen die Matura abgeschlossen haben (5,4%), sechs Testpersonen eine Fachhochschule oder Akademie besuchten (16,2%), 23 Teilnehmer eine Universität mit Magister abgeschlossen haben (62,2%) und drei getestete Personen ein Doktoratsstudium absolvierten (8,1%). Auch hier muss festgestellt werden, dass die Stichprobe vorwiegend aus Testpersonen mit höherer Ausbildung zusammengesetzt ist. Die Repräsentativität bezüglich dieses Kriteriums ist daher nicht gegeben.

Tabelle 19: Verteilung und Anteilswerte der höchsten abgeschlossenen Ausbildung im Wortflüssigkeitstest (N = 37); Angaben zur Kategorisierung siehe Tabelle 5

| Kategorie | Häufigkeit | Prozent |
|---------------|------------|------------|
| 0 | 3 | 8,1 |
| 1 | 0 | 0 |
| 2 | 0 | 0 |
| 3 | 2 | 5,4 |
| 4 | 6 | 16,2 |
| 5 | 23 | 62,2 |
| 6 | 3 | 8,1 |
| Gesamt | 37 | 100 |

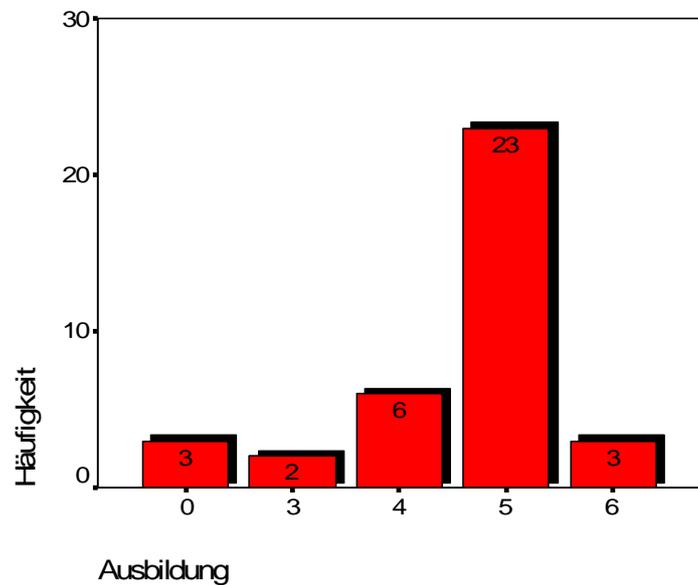


Abbildung 19: Verteilung der höchsten abgeschlossenen Ausbildung im Wortflüssigkeitstest (N = 37)

8.2. Reliabilität und Trennschärfen (N = 37)

Für den ursprünglichen Gesamttest (k= 20) würde sich ein Cronbach Alpha von 0,79 ergeben.

Durch Entfernung der negativ-trennscharfen Items (11 und 18) verbleiben schließlich 18

Items, die dann ein Cronbach Alpha von 0,82 und Trennschärfen zwischen

$r = 0,00$ und $r = 0,64$ aufweisen.

Die mit (*) gekennzeichneten Items (vgl. Tabelle 20) werden aufgrund negativer Trennschärfen eliminiert.

Tabelle 20: Itemkennwerte im Wortflüssigkeitstest (k= 20)

| Item | Lösungswahrscheinlichkeiten (p _i) | Standardabweichung (σ _i) | Korrigierte Itemtrennschärfe ρ (x _i , X*) | Alpha (if item deleted) |
|------|---|--------------------------------------|--|-------------------------|
| 1 | 0,27 | 0,45 | 0,54 | 0,77 |
| 2 | 0,35 | 0,48 | 0,39 | 0,78 |
| 3 | 0,38 | 0,49 | 0,60 | 0,76 |
| 4 | 0,14 | 0,35 | 0,47 | 0,77 |
| 5 | 0,54 | 0,50 | 0,55 | 0,76 |
| 6 | 0,92 | 0,28 | 0,02 | 0,79 |
| 7 | 0,46 | 0,51 | 0,49 | 0,77 |
| 8 | 0,54 | 0,50 | 0,51 | 0,77 |
| 9 | 0,86 | 0,35 | 0,30 | 0,78 |
| 10 | 0,92 | 0,28 | 0,45 | 0,78 |
| 11* | 0,11 | 0,31 | -0,11* | 0,80 |
| 12 | 0,89 | 0,31 | 0,53 | 0,77 |
| 13 | 0,30 | 0,46 | 0,47 | 0,77 |
| 14 | 0,19 | 0,40 | 0,57 | 0,77 |
| 15 | 0,19 | 0,40 | 0,53 | 0,77 |
| 16 | 0,59 | 0,50 | 0,41 | 0,77 |
| 17 | 0,78 | 0,42 | 0,15 | 0,79 |
| 18* | 0,43 | 0,50 | -0,11* | 0,81 |
| 19 | 0,32 | 0,47 | 0,19 | 0,79 |
| 20 | 0,65 | 0,48 | 0,13 | 0,79 |

8.3. Verteilung des Rohscores

Aufgrund der vorangegangenen Berechnungen (siehe Itemtrennschärfe, Itemschwierigkeiten) mussten zwei Items ausgeschlossen werden. Die hier berechnete Verteilung der Rohscores bezieht sich somit auf die verbleibenden 18 Items. Es kann ein minimaler Rohscore von 2 und ein maximaler Wert von 17 richtig gelösten Items abgelesen werden. Der Mittelwert (MW)

kann mit 9,30 bei einer Standardabweichung (SD) von 3,86 angegeben werden. Der Median liegt in dieser Stichprobe bei 9.

Tabelle 21: Deskriptivstatistische Kennzahlen zum Rohwert im Wortflüssigkeitstest (k = 18; N = 37)

| | Median | MW | SD | Minimum | Maximum |
|----------|--------|------|------|---------|---------|
| Rohscore | 9 | 9,30 | 3,86 | 2 | 17 |

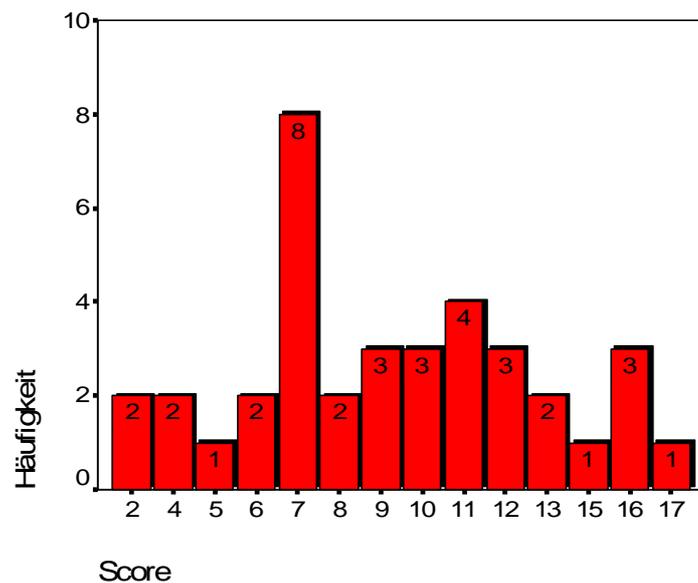


Abbildung 20: Verteilung des Rohscores im Wortflüssigkeitstest (k = 18, N = 37)

Die Verteilung des Rohscores im Wortflüssigkeits-Test, untersucht mittels Kolmogorov-Smirnov-Test ergibt $z = 0,79$ und $p = 0,57$. Die Normalverteilung der Daten kann somit angenommen werden.

8.4. Überprüfung auf Rasch- Homogenität

Von den ursprünglich $k = 20$ Items wurden 2 Items ausgeschlossen. Mit den 18 verbliebenen Items wird nun eine Überprüfung auf Rasch-Konformität durchgeführt. Der

Stichprobenumfang setzt sich aus jenen 37 Testpersonen zusammen, von denen anzunehmen ist, dass sie die Tests ernsthaft bearbeitet haben.

Tabelle 22: Schwierigkeitsparameter und Standardfehler der einzelnen Items im Wortflüssigkeitstest (Gesamtstichprobe)

| Item | Itemparameter | Standardfehler |
|-----------------|-----------------|----------------|
| 1 | +1,53 | 0,42 |
| 2 | +1,00 | 0,22 |
| 3 | +0,84 | 0,23 |
| 4 | +2,67 | 0,37 |
| 5 | -0,05 | 0,19 |
| 6 | -2,92 | 0,45 |
| 7 | +0,39 | 0,32 |
| 8 | -0,05 | 0,19 |
| 9 | -2,24 | 0,49 |
| 10 | -2,92 | 0,45 |
| 11 [*] | Nicht verwendet | |
| 12 | -2,55 | 0,54 |
| 13 | +1,34 | 0,30 |
| 14 | +2,16 | 0,34 |
| 15 | +2,16 | 0,34 |
| 16 | -0,35 | 0,25 |
| 17 | -1,52 | 0,41 |
| 18 [*] | Nicht verwendet | |
| 19 | +1,17 | 0,25 |
| 20 | -0,65 | 0,28 |

Anmerkung: Die Itemparameter in der Gesamtstichprobe wurden mit (-1) multipliziert.

Positive Werte sind daher als schwierigere Items zu verstehen, während negative Werte als leichtere Items gesehen werden können.

8.4.1. Teilungskriterium Geschlecht

Die Prüfung auf Rasch-Konformität der Items im Wortflüssigkeitstest-Test wird zunächst anhand des externen Teilungskriteriums Geschlecht (weiblich, männlich) durchgeführt.

Von den 37 gültigen Testpersonen sind 18 weiblich und 19 männlich.

Als globale Testgröße im Likelihood-Quotiententest wird Chi-Quadrat (χ^2) nach Andersen berechnet, die Schwierigkeitsparameter aus den Teilstichproben werden zur Prüfung der Modellgeltung für alle Items ($df = k - 1$) gemeinsam berücksichtigt (die Chi-Quadrat-Approximation wird nach Wilson-Hilferty durchgeführt)

Tabelle 23: Likelihood-Quotiententest mittels χ^2 nach Andersen im Wortflüssigkeitstest (Teilungskriterium Geschlecht)

| Teilungskriterium Geschlecht | |
|------------------------------|-------|
| Andersen - χ^2 | 16,04 |
| Freiheitsgrade (df) | 17 |
| χ^2 bei alpha 5% | 27,58 |
| χ^2 bei alpha 1% | 33,43 |

Die Prüfung fällt sowohl bei alpha = 5% als auch 1% nicht signifikant aus, die Modellgeltung nach dem Teilungskriterium Geschlecht kann zunächst als gegeben angenommen werden.

Die Überprüfung der Parameterschätzwertedifferenz auf Signifikanz ergibt allerdings pro Item, dass diese nicht schätzbar sind. Diese Beobachtung lässt darauf schließen, dass einerseits bestimmte Differenzen der Itemparameter beträchtlich sind und andererseits, dass der Stichprobenumfang in diesem Fall keine genaue Schätzung gestattet.

Auch weist der grafische Modelltest eher darauf hin, dass die Items relativ weit außerhalb der 45°-Geraden zu liegen kommen. Dieser Umstand kann als Diskrepanz zum nicht-signifikanten Ergebnis im Andersen- χ^2 -Test gewertet werden, sodass im externen

Teilungskriterium Geschlecht im Wortflüssigkeits-Test eine Verletzung der Rasch-Homogenität vorliegt.

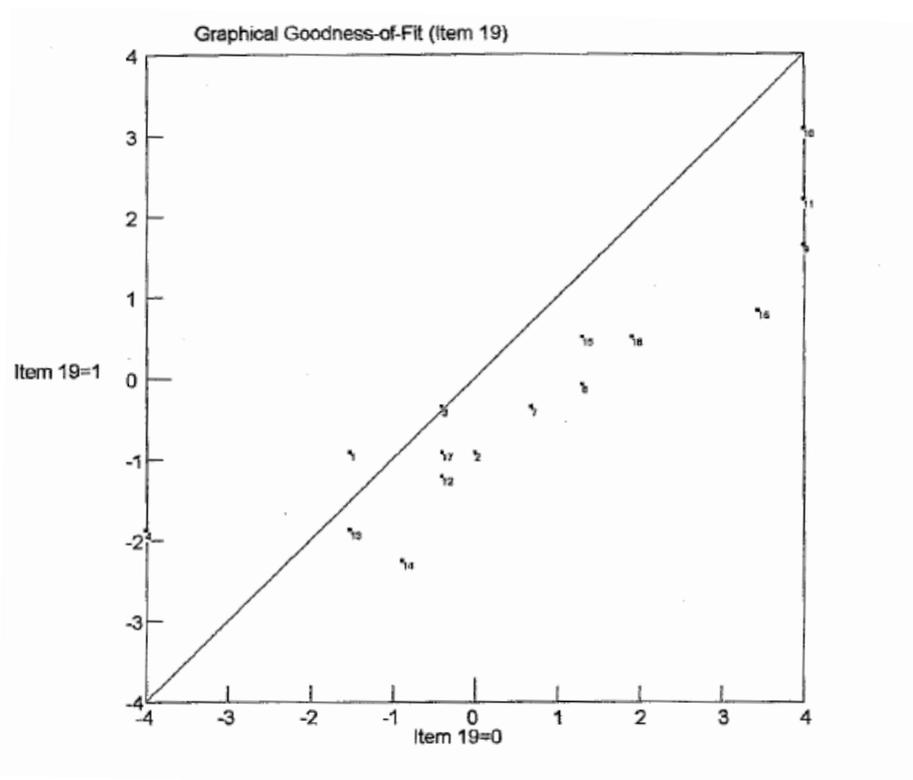


Abbildung 21: Grafische Modellprüfung im Wortflüssigkeitstest Teilungskriterium Geschlecht (k = 18; N = 37)

8.4.2. Teilungskriterium Alter

Als weiteres externes Teilungskriterium wird nun das Alter herangezogen. Der Cutting-Score der Stichprobe im Wortflüssigkeits-Test ist der Mittelwert des Lebensalters, Testpersonen bis 33 Jahre werden als „jung“ bezeichnet, jene mit einem höheren Lebensalter werden der Teilstichprobe „alt“ zugeordnet. 22 Personen (59,5%) können somit als „jung“, 15 Personen (40,5%) als „alt“ klassifiziert werden.

Tabelle 24: Likelihood-Quotiententest mittels χ^2 nach Andersen im Wortflüssigkeitstest (Teilungskriterium Alter)

| Teilungskriterium Alter | |
|-------------------------|-------|
| Andersen - χ^2 | 11,56 |
| Freiheitsgrade (df) | 17 |
| χ^2 bei alpha 5% | 27,58 |
| χ^2 bei alpha 1% | 33,43 |

Die Prüfung fällt sowohl bei alpha = 5% als auch 1% nicht signifikant aus, die Modellgeltung nach dem Teilungskriterium Alter kann als gegeben angenommen werden.

Die einzelnen Items werden nun im Detail zur Beurteilung der Signifikanz der Parameterschätzwertedifferenz mittels Wald-Statistik überprüft (Kritische z-Werte für alpha 5% = 1,96, für alpha 1% = 2,58).

Tabelle 25: Itemparameter der Teilstichproben („jung“ und „alt“) mit Signifikanzbeurteilung nach Wald im Wortflüssigkeitstest

| Item | Itemparameter „jung“ | Itemparameter „alt“ | z - Wert | Signifikanz |
|------|----------------------|---------------------|----------|-------------|
| 1 | 1,43 | 1,71 | 0,31 | n.s. |
| 2 | 1,43 | 0,39 | -1,80 | n.s. |
| 3 | 0,62 | 1,20 | 0,95 | n.s. |
| 4 | 2,41 | 3,65 | 1,09 | n.s. |
| 5 | 0,13 | -0,34 | -0,86 | n.s. |
| 6 | -3,44 | -2,64 | 0,85 | n.s. |
| 7 | 0,62 | 0,03 | -1,26 | n.s. |
| 8 | -0,12 | 0,03 | 0,27 | n.s. |
| 9 | -2,00 | -2,64 | -0,72 | n.s. |
| 10 | -2,58 | -3,53 | -0,72 | n.s. |
| 12 | -3,44 | -2,03 | 1,58 | n.s. |
| 13 | 1,43 | 1,20 | -0,40 | n.s. |
| 14 | 2,05 | 2,44 | 0,38 | n.s. |
| 15 | 1,73 | 3,65 | 1,88 | n.s. |
| 16 | 0,13 | -1,11 | -2,14 | sign. (5%) |
| 17 | -1,21 | -2,03 | -1,11 | n.s. |
| 19 | 1,43 | 0,77 | -1,03 | n.s. |
| 20 | -0,62 | -0,71 | -0,15 | n.s. |

Die Überprüfung der einzelnen Items ergibt, dass nur das Item 16 in diesen Teilstichproben unterschiedlich schwierig zu bezeichnen ist.

8.4.3. Teilungskriterium Rohscore

Als internes Teilungskriterium wird die Leistung im Test (niedriger vs. hoher Rohscore) selbst herangezogen.

Die Stichprobe wird aufgrund des Mittelwertes der Testleistung der Personen (MW = 9,30; Md = 9) geteilt: Die Teilstichproben setzen sich aus Personen mit „wenigen“ richtigen Antworten (Score ≤ 9) und jenen mit mehr korrekten Antworten (Score > 9) zusammen. Insofern haben 20 Personen (54,1%) wenige, 17 Personen (45,9%) mehr Items gelöst.

Tabelle 26: Likelihood-Quotiententest mittels χ^2 nach Andersen im Wortflüssigkeitstest (Teilungskriterium Rohscore)

| Internes Teilungskriterium Rohscore (MW) | |
|---|-------|
| Andersen - χ^2 | 19,63 |
| Freiheitsgrade (df) | 17 |
| χ^2 bei alpha 5% | 27,58 |
| χ^2 bei alpha 1% | 33.43 |

Die Prüfung fällt sowohl bei alpha = 5% als auch 1% nicht signifikant aus, die Modellgeltung nach dem Teilungskriterium Rohscore kann zunächst als gegeben angenommen werden. Die Überprüfung der Parameterschätzwertedifferenz auf Signifikanz ergibt allerdings pro Item, dass diese nicht schätzbar sind. Dieser Umstand lässt darauf schließen, dass einerseits die Differenzen der Itemparameter beträchtlich sind und andererseits, dass der Stichprobenumfang in diesem Fall erneut keine genaue Schätzung gestattet.

Der grafische Modelltest weist darauf hin, dass die Items weit außerhalb der 45°-Geraden zu liegen kommen. Diese Diskrepanz zum nicht-signifikanten Ergebnis im Andersen- χ^2 -Test deutet auf eine Verletzung der Rasch-Homogenität im internen Teilungskriterium im Wortflüssigkeitstest hin.

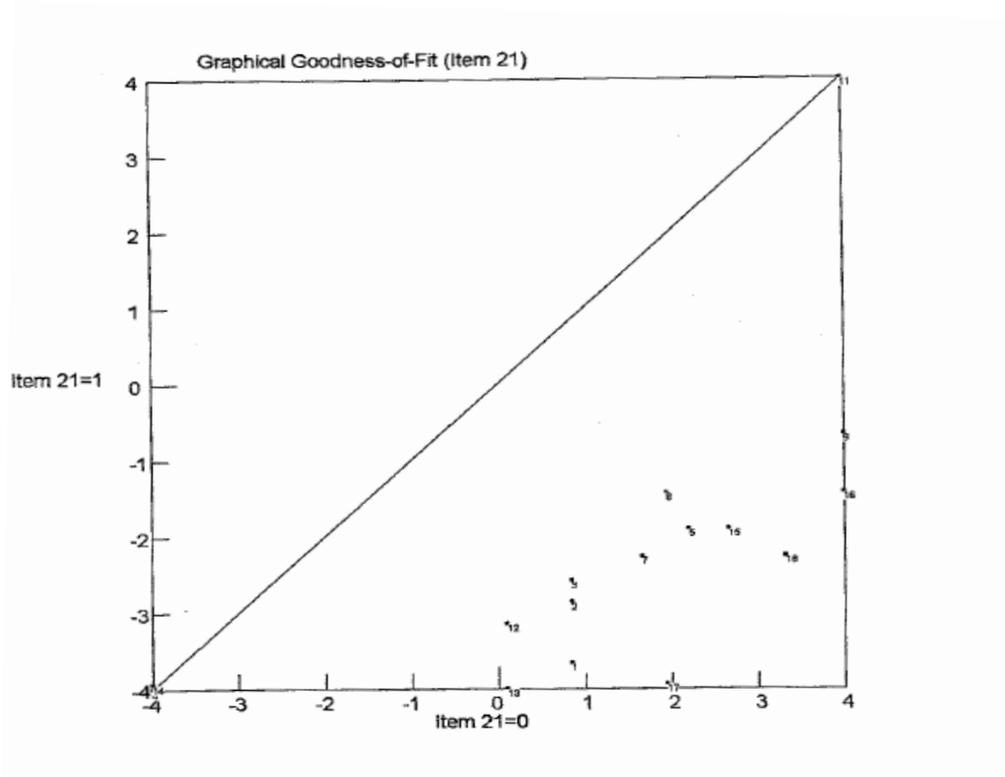


Abbildung 22: Grafische Modellprüfung im Wortflüssigkeitstest Teilungskriterium Rohscore (k = 18; N = 37)

8.4.4. Personenparameter im Wortflüssigkeitstest

Anhand der Datenanalyse geht im Sinne der erschöpfenden Statistiken hervor, dass Personen mit einer bestimmten Anzahl gelöster Items im Wortflüssigkeitstest einen eindeutigen Fähigkeitsparameter Θ erhalten.

Tabelle 27: Rohwerte und die entsprechenden Fähigkeitsparameter (Θ) im Wortflüssigkeitstest

| Rohwert | Theta (Θ) |
|---------|--------------------|
| 1 | -4,08 |
| 2 | -3,16 |
| 3 | -2,52 |
| 4 | -1,98 |
| 5 | -1,49 |
| 6 | -1,05 |
| 7 | -0,64 |
| 8 | -0,25 |
| 9 | +0,11 |
| 10 | +0,47 |
| 11 | +0,82 |
| 12 | +1,18 |
| 13 | +1,55 |
| 14 | +1,95 |
| 15 | +2,41 |
| 16 | +2,98 |
| 17 | +3,83 |

Im vorliegenden Wortflüssigkeitstest haben die Personen zwischen 2 und 17 Items gelöst. Der Fall, dass Personen keine oder alle Items gelöst haben, kommt in dieser Stichprobe nicht vor; d.h. die zugehörigen Parameter $\Theta_{r=0} = -\infty$ und $\Theta_{r=k} = +\infty$ sind nicht erforderlich.

9. Hypothesentestung

Die Grundlage der Hypothesentestungen bilden im Synonyme-Test 73 Personen, während im Wortflüssigkeitstest 37 Testpersonen untersucht wurden. Für die Berechnungen des Rohscores werden nur die selektierten Items herangezogen (SYN¹⁰: k = 14; WF¹¹: k = 18).

9.1. Zusammenhang der Leistungen in beiden Tests

H₁: Die Leistungen im Synonyme-Test und Wortflüssigkeitstest korrelieren positiv ($\rho > 0$).

Verfahren: Produkt-Moment-Korrelation nach Pearson

Voraussetzung: Normalverteilung und zumindest intervallskalierte Daten, die in einem linearen Zusammenhang stehen.

Zur Berechnung der Produkt-Moment-Korrelation wurden jene 27 Testpersonen herangezogen, die beide Tests vollständig bearbeitet haben.

Die Überprüfung mittels Kolmogorov-Smirnov-Anpassungstest ergibt, wie in Tabelle 28 ersichtlich ist, dass aufgrund der nicht signifikanten Ergebnisse eine Normalverteilung der Daten angenommen werden kann.

¹⁰ Synonyme-Test (Synonyme Fragebogen engl (Version 1)

¹¹ Wortflüssigkeitstest (Word Fluency Test „WF engl“)

Tabelle 28: Beurteilung der Normalverteilung der Daten im Synonyme-Test (SYN) und Wortflüssigkeitstest (WF)

| | SYN | WF |
|--------------|--------------|--------------|
| K-S z | 1,06 | 0,64 |
| p | 0,211 (n.s.) | 0,807 (n.s.) |

Tabelle 29: Deskriptivstatistische Kennwerte zum Synonyme-Test (SYN) und Wortflüssigkeitstest (WF)

| | MW | SD | N |
|------------|-----------|-----------|----------|
| SYN | 8,41 | 2,17 | 27 |
| WF | 9,30 | 4,03 | 27 |

Der Korrelationskoeffizient $r = -0,16$ fällt mit $p = 0,414$ nicht signifikant aus, d.h. die beiden Testteile können als unabhängig voneinander angesehen werden. Der erklärte Varianzanteil fällt mit $r^2 = 0,026$ sehr gering aus.

Die Hypothese, dass beide Tests miteinander korrelieren muss daher zurückgewiesen werden.

9.2. Varianzanalyse - Synonyme Fragebogen engl (Version 1)

$H_1^{(1)}$: Es gibt einen Unterschied in der Leistung bei Synonyme finden bezüglich der Alterskategorie.

$H_1^{(2)}$: Es gibt einen Unterschied in der Leistung bei Synonyme finden bezüglich des Geschlechts.

$H_1^{(3)}$: Es gibt eine Wechselwirkung zwischen Geschlecht und Alter.

Verfahren: Zweifaktorielle, univariate Varianzanalyse.

Voraussetzung: Normalverteilung und zumindest Intervallskalierung der Daten; Homogenität der Varianzen

UV: Geschlecht, Alter (jung, alt)

AV: Score im Synonyme-Test

Die Überprüfung der Normalverteilung der Daten mittels Kolmogorov-Smirnov-Anpassungstest ergibt, wie in Tabelle 30 zu sehen ist, folgende Ergebnisse.

Tabelle 30: Beurteilung der Normalverteilung der Daten in den vier Faktorstufenkombinationen im Synonyme-Test

| | weiblich-jung | weiblich-alt | männlich-jung | männlich-alt |
|--------------|----------------------|---------------------|----------------------|---------------------|
| K-S z | 0,95 | 0,57 | 0,71 | 0,78 |
| p | 0,334 (n.s.) | 0,902 (n.s.) | 0,702 (n.s.) | 0,571 (n.s.) |

Jeweils kann aufgrund der nicht signifikanten Resultate eine Normalverteilung der Daten angenommen werden.

Die Überprüfung der Homogenität der Varianzen mittels Levene-Test ergibt $F = 0,50$ ($df_1 = 3$, $df_2 = 69$), $p = 0,680$. Die Nullhypothese, dass die Fehlervarianz der abhängigen Variablen über Gruppen hinweg gleich ist, kann beibehalten werden.

Tabelle 31: Deskriptivstatistik zum Synonyme-Test in Abhängigkeit von Alter und Geschlecht

| Alterskategorie | Geschlecht | MW | SD | N |
|-----------------|---------------|-------------|-------------|-----------|
| jung | w | 8,04 | 1,79 | 25 |
| | m | 7,94 | 2,24 | 16 |
| | Gesamt | 8,00 | 1,95 | 41 |
| alt | w | 9,29 | 1,69 | 17 |
| | m | 9,60 | 1,88 | 15 |
| | Gesamt | 9,44 | 1,76 | 32 |
| Gesamt | w | 8,55 | 1,84 | 42 |
| | m | 8,74 | 2,21 | 31 |
| | Gesamt | 8,63 | 1,99 | 73 |

Die Berechnung einer zweifaktoriellen, univariaten Varianzanalyse ergibt folgende Resultate (vgl. Tabelle 32)

Tabelle 32: Tafel der Varianzanalyse im Synonyme-Test

| | Abweichungs- Quadratsumme (AQS) | df | Varianz- schätzung (σ^2) | F | Sign. (p) | Effektgröße (η^2) |
|-------------------------------|---------------------------------------|----|---|-------|-----------|-----------------------------|
| Alter | 37,31 | 1 | 37,31 | 10,42 | 0,002 ** | 0,131 |
| Geschlecht | 0,18 | 1 | 0,18 | 0,05 | 0,823 | 0,001 |
| Alter x Geschlecht | 0,73 | 1 | 0,73 | 0,20 | 0,653 | 0,003 |
| Fehler | 247,03 | 69 | 3,58 | | | |
| Gesamt (korr.) | 285,25 | 72 | | | | |

Die $H_1^{(1)}$, dass es einen Unterschied in der Leistung bei Synonyme finden bezüglich der Alterskategorie (jung vs. alt) gibt, kann aufgrund des signifikanten Ergebnisses übernommen werden (vgl. Abbildung 23). Ältere Personen weisen hier höhere Werte auf. Die

entsprechende partielle Effektgröße ($\eta^2 = 0,131$) weist darauf hin, dass mit diesem Ergebnis der gegebene Unterschied in den Populationen unter Annahme eines mittelgroßen Effekts als relevant bezeichnet werden kann (vgl. Bortz, 2006, S. 606).

Die $H_1^{(2)}$, dass es einen Unterschied bezüglich des Geschlechts gibt, muss zurückgewiesen werden.

Die $H_1^{(3)}$, dass es Interaktionen zwischen Alter und Geschlecht gibt, muss ebenfalls zurückgewiesen werden. Dieses nicht-signifikante Ergebnis zur Wechselwirkung gestattet, die signifikante Hauptwirkung im Faktor Alter ohne Einschränkung interpretieren zu können.

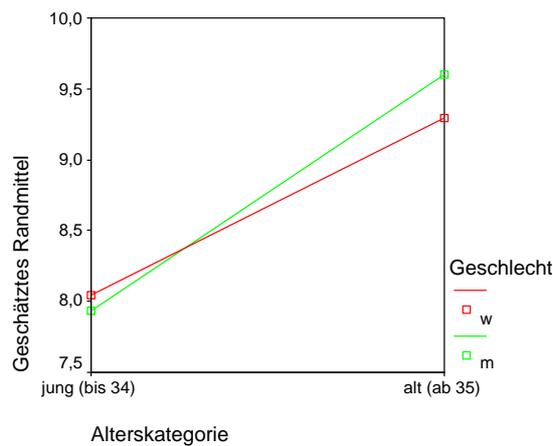


Abbildung 23: Mittelwerte im Synonyme-Test in Abhängigkeit von Geschlecht und Alterskategorie

9.3. Varianzanalyse - Word Fluency Test „WF engl“

$H_1^{(1)}$: Es gibt einen Unterschied in der Leistung bei Wortflüssigkeits-Aufgaben bezüglich der Alterskategorie.

$H_1^{(2)}$: Es gibt einen Unterschied in der Leistung bei Wortflüssigkeitsaufgaben im Geschlecht.

$H_1^{(3)}$: Es gibt eine Wechselwirkung zwischen Alter und Geschlecht.

Verfahren: Zweifaktorielle, univariate Varianzanalyse

Voraussetzung: Normalverteilung und zumindest Intervallskalierung der Daten;

Homogenität der Varianzen

UV: Geschlecht, Alter (jung, alt)

AV: Score im Wortflüssigkeitstest

Die Überprüfung der Normalverteilung der Daten mittels Kolmogorov-Smirnov-Anpassungstest ergibt wie in Tabelle 33 ersichtlich, folgende Ergebnisse.

Tabelle 33: Beurteilung der Normalverteilung der Daten in den vier Faktorstufenkombinationen im Wortflüssigkeitstest

| | weiblich-jung | weiblich-alt | männlich-jung | männlich-alt |
|--------------|---------------|--------------|---------------|--------------|
| K-S z | 1,03 | 0,52 | 0,43 | 0,48 |
| p | 0,240 (n.s.) | 0,950 (n.s.) | 0,993 (n.s.) | 0,975 (n.s.) |

Aufgrund der nicht signifikanten Resultate kann jeweils eine Normalverteilung der Daten angenommen werden.

Die Überprüfung der Homogenität der Varianzen mittels Levene-Test ergibt $F = 0,38$ ($df_1 = 3$, $df_2 = 33$), $p = 0,771$. Die Nullhypothese, dass die Fehlervarianz der abhängigen Variablen über Gruppen hinweg gleich ist, kann beibehalten werden.

Tabelle 34: Deskriptivstatistik zum Wortflüssigkeitstest in Abhängigkeit von Alter und Geschlecht

| Alterskategorie | Geschlecht | MW | SD | N |
|-----------------|---------------|-------------|-------------|-----------|
| jung | w | 8,92 | 3,40 | 12 |
| | m | 10,60 | 4,50 | 10 |
| | Gesamt | 9,68 | 3,93 | 22 |
| alt | w | 7,00 | 3,58 | 6 |
| | m | 9,89 | 3,72 | 9 |
| | Gesamt | 8,73 | 3,83 | 15 |
| Gesamt | w | 8,28 | 3,48 | 18 |
| | m | 10,26 | 4,05 | 19 |
| | Gesamt | 9,30 | 3,86 | 37 |

Die Berechnung einer zweifaktoriellen, univariaten Varianzanalyse ergibt folgende Resultate (vgl. Tabelle 35).

Tabelle 35: Tafel der Varianzanalyse im Wortflüssigkeitstest

| | Abweichungs- Quadratsumme (AQS) | df | Varianz- schätzung (σ^2) | F | Sign. (p) | Effektgröße (η^2) |
|-------------------------------|---------------------------------------|----|---|------|-----------|-----------------------------|
| Alter | 14,98 | 1 | 14,98 | 1,02 | 0,320 | 0,030 |
| Geschlecht | 45,34 | 1 | 45,34 | 3,09 | 0,088 | 0,086 |
| Alter x Geschlecht | 3,15 | 1 | 3,15 | 0,22 | 0,646 | 0,006 |
| Fehler | 484,21 | 33 | 14,67 | | | |
| Gesamt (korr.) | 547,68 | 36 | | | | |

Die $H_1^{(1)}$, dass es einen Unterschied in der Leistung bei Wortflüssigkeit bezüglich der Alterskategorie (jung vs. alt) gibt, muss aufgrund des nicht-signifikanten Ergebnisses zurückgewiesen werden (vgl. Abbildung 24).

Die $H_1^{(2)}$, dass es einen Unterschied bezüglich des Geschlechts gibt, wird zunächst zurückgewiesen, wobei jedoch Männer tendenziell höhere Werte erreichen.

Die $H_1^{(3)}$, dass es Interaktionen zwischen Alter und Geschlecht gibt, muss ebenfalls zurückgewiesen werden; es werden somit keine Wechselwirkungen angenommen.

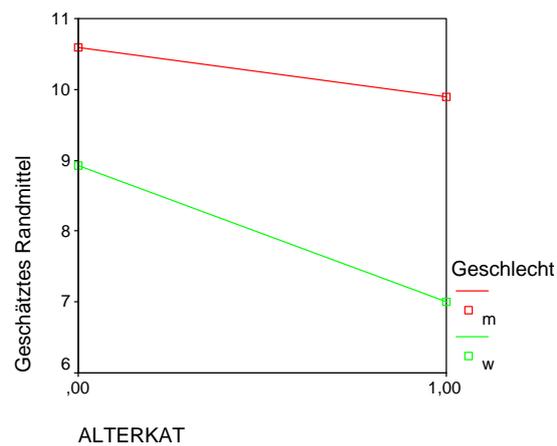


Abbildung 24: Mittelwerte im Wortflüssigkeitstest in Abhängigkeit von Geschlecht und Alterskategorie

10. Interpretation der Ergebnisse und Diskussion

Mit diesen beiden Verfahren wurden letztendlich 73 Personen (native speakers) im Synonyme-Test und 37 Personen im Wortflüssigkeitstest auf ihre verbalen Fähigkeiten getestet. Der geringe Stichprobenumfang, vor allem im Wortflüssigkeitstest, ist vermutlich darauf zurückzuführen, dass zur Testteilnahme Englisch als Muttersprache Voraussetzung war und dass außerdem die Instruktionen im TestWeb selbst nur auf Deutsch zur Verfügung standen. Trotz der englischsprachigen Anleitung, die den Testpersonen zusätzlich mitgeschickt wurde, war es für manche Personen schwierig die Anweisungen zu verstehen. Weiters ist anzumerken, dass der Wortflüssigkeitstest in der Reihenfolge als zweiter angeboten wurde, wodurch wahrscheinlich die Motivation nach dem ersten bearbeiteten Test nicht mehr so groß war. Nachdem ausschließlich Personen mit Englisch als Muttersprache an der Studie teilnehmen sollten, konnte das Internet als Medium optimal genutzt werden, um auch Personen im Ausland zu erreichen. Als großer Nachteil muss hier allerdings angeführt werden, dass man als Testleiter keinen Einfluss auf die Rahmenbedingen (Standardisierung der Testsituation, Störvariablen usw.) hat. Es kann daher nicht garantiert werden, dass die Personen die Verfahren instruktionsgemäß bearbeitet haben.

Zu den vorgegebenen Verfahren kann gesagt werden, dass die ursprünglichen Testlängen von jeweils 20 Items aufgrund der beobachteten Itemtrennschärfen und Itemschwierigkeiten auf $k_{\text{syn}} = 14$ und $k_{\text{wf}} = 18$ reduziert wurden. Diese verkürzten Tests zeigen unterschiedlich hohe Reliabilitäten. Die untere Schranke der Reliabilität nach Cronbach's Alpha beträgt für den Synonyme-Test 0,42, was eine niedrige Konsistenz der Items vermuten lässt und somit auf eine geringe Zuverlässigkeit hinweist. Im Gegensatz dazu kann im Wortflüssigkeitstest ein befriedigend hohes Cronbach-Alpha mit 0,82 festgestellt werden.

Anhand der Analysen der beiden Testverfahren auf Rasch-Homogenität ergibt sich folgendes Bild. Für beide Tests wurden als externe Teilungskriterien Geschlecht und Alter sowie als internes Teilungskriterium der Rohscore herangezogen. Das geplante externe Teilungskriterium Bildung konnte aufgrund der Überrepräsentation höher gebildeter Personen in der Stichprobe nicht verwendet werden. Für den Synonyme-Test ergeben sich für die externen Teilungskriterien bei globaler Betrachtung keinerlei Einschränkungen bezüglich der Rasch-Homogenität. Dies bedeutet, dass hier die gleiche latente Eigenschaftsdimension in fairer Weise gemessen wird. Bezüglich des internen Teilungskriteriums zeigen sich allerdings Diskrepanzen zwischen der globalen Modellgüte (Andersen- χ^2) und dem grafischen Modelltest, wobei die z-Werte der einzelnen Items nicht schätzbar waren. Für verschieden fähige Personen misst der Synonyme-Test offenbar nicht fair.

Im Wortflüssigkeitstest kann für das externe Teilungskriterium Alter die Rasch-Homogenität angenommen werden. Bezüglich des externen Teilungskriteriums Geschlecht sowie des internen Teilungskriteriums Rohscore zeigen sich erneut gewisse Abweichungen zwischen der globalen Modellgüte (Andersen- χ^2) und dem grafischen Modelltest, wobei die z-Werte der einzelnen Items hier ebenfalls nicht schätzbar waren. Für Personen unterschiedlichen Geschlechts sowie für verschieden fähige Testpersonen ist der Wortflüssigkeitstest offenbar nicht fair. Den beobachteten Widersprüchlichkeiten liegt möglicherweise eine zu niedrige Stichprobengröße zugrunde. Um voreilige Schlüsse zu vermeiden, müssen daher die nicht-signifikanten Ergebnis im Andersen- χ^2 -Test mit Vorbehalt betrachtet werden.

Die Korrelation der beiden Tests, berechnet auf Grundlage jener 27 Personen, die beide Verfahren vollständig bearbeitet haben, lässt eher auf eine Unabhängigkeit der beiden Verfahren schließen. Die beiden Tests messen offenbar nicht dasselbe, d.h. Wortflüssigkeit und Synonyme finden können insofern als unterschiedliche Dimensionen betrachtet werden.

Diese Ergebnisse gehen konform mit den Erkenntnissen von Thurstone, der in seiner Theorie der „primary mental abilities“ unter anderem diese zwei nebeneinander stehende Faktoren (verbal comprehension und word fluency) postuliert.

In Anbetracht der formulierten Hypothesen, ob sich die Personen hinsichtlich ihres Geschlechts und der Alterskategorie bei den Testleistungen voneinander unterscheiden, können im Synonyme-Test ältere Personen als fähiger eingestuft werden, während Frauen und Männer keinen Unterschied aufweisen. Wie auch schon in Kapitel fünf beschrieben wurde, geht auch Wechsler von der Annahme aus, dass der Wortschatz relativ altersunabhängig ist. Das Identifizieren von Synonymen scheint nach Kohrt und Kucharczik (1998) bis ins hohe Alter recht stabil zu sein. Diese Annahmen bezüglich des Alters können in der vorliegenden Stichprobe recht gut beobachtet werden.

Im Wortflüssigkeitstest treten keine signifikanten Unterschiede zwischen den Alterskategorien und dem Geschlecht auf, allerdings erbringen Männer hier überraschenderweise tendenziell bessere Leistungen. Laut Halpern (2000) beschreiben unterschiedliche Quellen eine durchschnittlich bessere Leistung der Frauen im Hinblick auf verbalen Fähigkeiten. Hyde und Linn (1998) weisen jedoch darauf hin, dass die Effektgrößen in unterschiedlichen verbalen Fähigkeitstests nur sehr gering ausfallen, wodurch praktisch von keinem Geschlechtsunterschied gesprochen werden kann. Weiters berichtet Light (1993, zitiert nach Kohrt & Kucharczik, 1998) von Wortfindungsschwierigkeiten älterer Personen, die hier in der vorliegenden Stichprobe offensichtlich nicht zu tragen kommen. Eine mögliche Erklärung dafür ist, dass diese genannten Schwierigkeiten wahrscheinlich erst in einem eher hohen Alter auftreten und die getesteten Personen in dieser Studie relativ jung waren.

11. Zusammenfassung

Hinsichtlich der Evaluierung der beiden Sprachtests zur Erfassung der Wortflüssigkeit (WF engl) und des Wortschatzes (Synonyme Fragebogen engl-Version 1) wurde im Zeitraum vom 24. April und 9. Juli 2008 eine Online-Studie anhand englischsprachiger Testpersonen durchgeführt. Von den ursprünglich 83 Personen im Synonyme-Test und den 42 Personen im Wortflüssigkeitstest blieben aufgrund unglaublicher Testbearbeitungszeiten letztendlich 73 Personen im ersteren Test und 37 Probanden im letzteren Verfahren übrig. Weiters wurden die beiden Tests um jene Items verkürzt, die aufgrund der Trennschärfen und Itemschwierigkeiten unpassend erschienen. Von den jeweils 20 ursprünglichen Items, mussten im Synonyme-Test die Items 1, 12, 14, 16, 17 und 19 entfernt werden. Im Wortflüssigkeitstest hingegen wurden lediglich die Items 11 und 18 ausgeschieden. Auf Basis dieser korrigierten Daten wurde eine relativ niedrige Reliabilität im Synonyme-Test festgestellt. Die Zuverlässigkeit im Wortflüssigkeitstest kann jedoch als zufriedenstellend betrachtet werden.

Die Rasch-Homogenität kann im Synonyme-Test nur in den externen Teilungskriterien (Geschlecht und Alter) uneingeschränkt angenommen werden. Im internen Teilungskriterium (Rohscore) kommt es zu uneindeutigen Ergebnissen aufgrund von Diskrepanzen zwischen der globalen Modellgüte (Andersen- χ^2) und der grafischen Modellprüfung, sowie den nicht schätzbaren z-Werten. Für den Wortflüssigkeitstest, kann hingegen nur die Rasch-Homogenität für das Teilungskriterium Alter angenommen werden. In den Teilungskriterien Geschlecht und Rohscore ergeben sich ebenfalls Widersprüche, wie schon beim Synonyme-Test bereits beschrieben wurden.

Berechnungen zur Korrelation zwischen beiden Tests konnten keinen Zusammenhang feststellen.

Im Synonyme-Test stellte sich heraus, dass ältere Personen bessere Leistungen erzielten, Geschlechtsunterschiede konnten allerdings nicht nachgewiesen werden.

Bezüglich des Wortflüssigkeitstest ergaben sich weder Alters- noch Geschlechtsunterschiede.

Es kann lediglich beobachtet werden, dass Männer tendenziell bessere Leistungen erbringen.

LITERATURVERZEICHNIS

Acker B. (2001). Sprachverständnis und Sprachbegabung. Von der Analyse Vorhandener Verfahren zu einem neuen Test. Landau: Verlag Empirische Pädagogik.

Aitichson J. (1997). *Wörter im Kopf. Eine Einführung in das mentale Lexikon*. Tübingen: Max Niemeyer Verlag.

Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2006). *Differentielle Psychologie und Persönlichkeitsforschung*. (6. vollständig überarbeitete Aufl.). Stuttgart: Kohlhammer.

Anschütz, G. & Ruttman, W. J. (1912). *Alfred Binet – Die neuen Gedanken über das Schulkind*. Leipzig: Verlag von Ernst Wunderlich.

Aschenbrenner, S., Tucha, O. & Lange, K. W. (2000). RWT: Regensburger Wortflüssigkeits-Test (Handanweisung). Göttingen: Hogrefe

Blutner R. (1986). Lexikalischer Zugriff: Experimentelle Paradigmen und Erklärungsansätze. In R. Blutner & K. Goede (Hrsg.), *Mentales Lexikon – Zugriff, Verarbeitung, Erwerb*. (S. 1-23). VEB Kongreß- und Werbeindruck Oberlungwitz

Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation – für Human- und Sozialwissenschaftler*. (4. überarbeitete Auflage). Heidelberg: Springer Verlag.

Brunnschweiler, Th. (2004). Magie, Manie, Manier. Versuch über die Geschichte des Anagramms. In M. Ch. Graeff (Hrsg.) *Die Welt hinter den Wörtern. Die Geschichte und Gegenwart des Anagramms* (S. 17-86). Alpnach: Verlag Martin Wallimann

Carroll, J. B. (1993). *Human cognitive abilities: a survey of the factor-analytic studies*. Cambridge: Cambridge University Press.

Clahsen, H. (1982). *Spracherwerb in der Kindheit. Eine Untersuchung zur Entwicklung der Syntax bei Kleinkindern*. Tübingen: Gunter Narr Verlag.

Dittmann, J. (2002). *Der Spracherwerb des Kindes. Verlauf und Störungen*. München: Verlag C. H. Beck.

Engelkamp, J. (1995). Mentales Lexikon: Struktur und Zugriff. In G. Harras (Hrsg.), *Die Ordnung der Wörter. Kognitive und lexikalische Strukturen*. (S. 99-119). Berlin, New York: Walter de Gruyter.

Fröhlich, W. D. (2002). *Wörterbuch Psychologie*. München: Deutscher Taschenbuch Verlag.

Funke, J. & Vaterrodt-Plünnecke, B. (1998). *Was ist Intelligenz?* München: Verlag C.H. Beck

Funke, J. (2006). Alfred Binet (1857-1911) und der erste Intelligenztest der Welt. In G. Lamberti (Hrsg.), *Intelligenz auf dem Prüfstand. 100 Jahre Psychometrie* (S. 22-39). Göttingen: Vandenhoeck & Ruprecht.

Gallistl, J. (2006). *Unterschiede in der Wortflüssigkeit – Auswirkungen von Geschlecht, Bildung und Alter*. Unveröffentlichte Diplomarbeit, Universität Wien.

Gardner, H. (1991). *Abschied vom IQ: Die Rahmentheorie der vielfachen Intelligenzen*. (1. Auflage) (Heim, M. Übers.). Stuttgart: Klett-Cotta. (Original erschienen 1985: *Frames of Mind. The Theory of multiple Intelligences*)

Gardner, H. (2002). *Intelligenzen: Die Vielfalt des menschlichen Geistes* (1. Aufl.) (Spengler, U., Übers.). Stuttgart: Klett-Cotta. (Original erschienen 1999: *Intelligences Reframed. Multiple Intelligences for the 21st Century*)

Gittler, G. & Arendasy, M. (2000) *Differentielle Psychologie I – Grundlagen und Intelligenztheorien* (2. erweiterte Auflage). Wien: Eigenverlag (Skriptum)

Glück, H. (Hrsg.). (2005). *Metzler Lexikon Sprache*. (3., neuüberarbeitete Aufl.). Stuttgart, Weimar: Verlag J.B. Metzler

Guilford, J. P. & Hoepfner, R. (1976). *Analyse der Intelligenz*. Weinheim, Basel: Beltz Verlag.

Halpern, D. F. (2000). *Sex Differences In Cognitive Abilities*. (Third Edition). Mahawa, NJ: Lawrence Erlbaum .

Herkner, W. (2001). *Lehrbuch Sozialpsychologie*. (2. unveränderte Aufl.). Bern: Verlag Hans Huber.

Holling, H., Preckel, F. & Vock M. (2004). *Psychologische Diagnostik*. (Band 6). Göttingen: Hogrefe.

Horn, W. (1962). Leistungsprüfsystem L.P.S.: Handanweisung für die Durchführung, Auswertung und Interpretation. Göttingen: Hogrefe

Huber, J.K. (2008). *Differentialpsychologische Aspekte verbaler Intelligenz: Synonyme*. Unveröffentlichte Diplomarbeit, Universität Wien.

Hyde, J. & Linn, M. (1988). Gender Differences in Verbal Ability: A Meta-Analysis. *Psychological Bulletin*, 104, 53-69.

Jäger, A. O. (1973). *Dimensionen der Intelligenz*. (3. Auflage). Göttingen: Hogrefe

Jescheniak, J. D. (2002). *Sprachproduktion. Der Zugriff auf das lexikale Gedächtnis beim Sprechen*. Göttingen: Hogrefe Verlag.

Kohrt, M. & Kucharczik, K. (1998). Sprache – unter besonderer Berücksichtigung von Jugend und Alter. In Fiehler, R. & Thimm, C. (Hrsg.) *Sprache und Kommunikation im Alter* (S. 17-37). Opladen / Wiesbaden: Westdeutscher Verlag.

Kubinger, K.D. (2006). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe Verlag

Langenmayr, A. (1997). *Sprachpsychologie. Ein Lehrbuch*. Göttingen: Hogrefe.

Meibauer, J. & Rothweiler, M. (Hrsg.). (1999). *Das Lexikon im Spracherwerb*. Tübingen, Basel: Francke Verlag

Meili, R. (1981). *Struktur der Intelligenz. Faktorenanalytische und denkpsychologische Untersuchungen*. Bern: Verlag Hans Huber.

Miller, G. A. (1993). *Wörter - Streifzüge durch die Psycholinguistik*. Heidelberg: Spektrum Akademischer Verlag. (Original erschienen 1991: *The Science of Words*).

Okupski, K. (2006). *Evaluation eines automatisch generierten Tests zur Wortflüssigkeit*. Unveröffentlichte Diplomarbeit, Universität Wien.

Pflaum, E. F. (2000). Les idées modernes sur les enfants (1909). In H. E. Lück, R. Miller & G. Sewz-Vosshenrich (Hrsg.), *Klassiker der Psychologie*. (S. 90–99). Stuttgart: Kohlhammer.

Probst, E. (1960). *Der Binet-Simon-Test zur Prüfung der Intelligenz bei Kindern*. (5. verbesserte und ergänzte Aufl.). Basel, New York: Karger.

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. (2., vollständig überarbeitete und erweiterte Auflage). Bern: Verlag Hans Huber.

Roth, E. (Hrsg.). (1998). *Intelligenz: Grundlage und neuere Forschung*. Stuttgart: Kohlhammer

Schuhfried GmbH, [Online] Url: <http://www.schuhfried.at/index.php?id=389&L=4> (Stand: 10. September 2008).

Schwibbe, G. (1984). *Intelligenz und Sprache: Zur Vorhersagbarkeit des intellektuellen Niveaus mittels kontentanalytischer Indikatoren*. Bochum: Brockmeyer.

Strube, G. (1977). Einleitung zu Band V. In G. Strube (Hrsg.), *Die Psychologie des 20. Jahrhunderts*. (Band 4, Binet und die Folgen); (S. 1-7). Zürich: Kindler Verlag.

Süllwold, F. (1977). Intelligenzdiagnostik und Intelligenztheorie. In G. Strube (Hrsg.), *Die Psychologie des 20. Jahrhunderts*. (Band 4, Binet und die Folgen); (S. 262-285). Zürich: Kindler Verlag.

Süß, H. M. (2003). Intelligenztheorien. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik*. (S. 217-224). Berlin: Beltz.

LEBENS LAUF

- Persönliche Daten:** Petra Laa
Geburtsdatum: 25. Mai 1982
Geburtsort: Wien
österreichische Staatsbürgerschaft
Religionsbekenntnis: röm.-kath.
Familienstand: ledig
- Familie:** Irene Laa, Bankangestellte
Ing. Manfred Laa, Pilot, †
Geschwister: 1 Schwester, 27 Jahre, Bankangestellte
- Schulbildung:** 1988 – 1992 Volksschule, Rannersdorf
1992 – 2000 BRG Schwechat
Matura Juni 2000
Seit Oktober 2000 Universität Wien (Psychologie)
Wahlfächer: Pädagogik, Klinische Psychologie,
Sozialpsychologie
Diplomarbeit im Bereich der Psychologischen Diagnostik
- Sonstige Ausbildung:** 11.11.1997 – vom Team zu Team – perfect training
25.02.1999 – Lernen ohne Leiden – perfect training
September 2001 bis Juni 2002 – Gebärdensprachkurs
(WITAF)
- Arbeitserfahrung:** Mai 1998 – MC Donald´s
August 1998 – Pensionsversicherungs-Anstalt
September 2000 bis Februar 2001 – OGM
(Meinungsforschungsinstitut)
November bis Dezember 2000 – H&M
März 2001 bis Oktober 2005 – Hervis
Seit August 2005 – HTM Sport- und Freizeitgeräte AG
- Praktikumserfahrung:** September bis Oktober 2005 - Beratungszentrum für
Schulfragen

Schwechat, am 10. September 2008