

TEMEKKO

Validierung eines Tests zur Messung exekutiver Kontrolle / Konzentration

Verena Schmid



München 2010

TEMEKKO

Validierung eines Tests zur Messung exekutiver Kontrolle / Konzentration

Inaugural - Dissertation

zur Erlangung des Doktorgrades der Philosophie an der

Fakultät für Psychologie und Pädagogik der

Ludwig-Maximilians-Universität, München

vorgelegt von

Verena Schmid

März 2010

Betreuer der Dissertation:

Referent:

Prof. Dr. Markus Bühner

Koreferent:

Prof. Dr. Matthias Ziegler

Tag der mündlichen Prüfung:

30. April 2010

**Messen, was messbar ist,
und messbar machen,
was noch nicht messbar ist.**

(Galileo Galilei 1564 - 1642)

Vorwort

Im Zuge der Bearbeitung meines Themas habe ich viel Unterstützung erfahren. Ganz herzlich bedanke ich mich an dieser Stelle bei all jenen, die zum Entstehen und zum Gelingen der Arbeit beigetragen haben.

An erster Stelle gilt mein Dank meinem Doktorvater, Prof. Dr. Markus Bühner. Die freundschaftliche Zusammenarbeit, direkte Ansprechbarkeit sowie unsere zahlreichen regelmäßigen Treffen erlaubten es, Gedanken und Ideen konstruktiv zu diskutieren und zielorientiert Lösungsansätze zu erarbeiten. Inspiration und Motivation prägten den Betreuungsprozess auf ganz besondere Weise und ermöglichten es mir, mich fachlich und persönlich weiter zu entwickeln, viel zu lernen und wertvolle Erfahrungen zu machen.

Ferner bedanke ich mich bei Prof. Dr. Matthias Ziegler für die Bereitschaft als Koreferent zu fungieren, die vielen kostbaren Hinweise sowie für die wertvolle Unterstützung an wichtigen Zwischenschritten im Bearbeitungsprozess.

Herzlichen Dank!

Im Sinne einer besseren Lesbarkeit beschränke ich mich in meiner Arbeit sprachlich auf die männliche Form von Personen bezogenen Substantiven. Dies impliziert keinesfalls eine Benachteiligung weiblicher Personen. Wenn beispielsweise vom Versuchsleiter oder vom Probanden gesprochen wird, so beinhaltet dies selbstverständlich auch die Versuchsleiterin oder die Probandin.

Zusammenfassung

Ziel dieser Untersuchungen war es, einen neu entwickelten Test zu validieren, der vorgab, Konzentration bzw. exekutive Kontrolle zu erfassen. Dieses Verfahren, der TEMEKKO, wurde nach dem top-down Ansatz entwickelt. Anhand einer Sammlung von in der Literatur wiederkehrenden Merkmalen der Konzentration sowie einer zu Grunde gelegten Definition von Konzentration wurden nach rationaler Methode die Items als objektive Indikatoren zur Erfassung der individuellen Fähigkeitsausprägung konstruiert.

Der TEMEKKO wurde als Durchstreichetest konzipiert. Dem Probanden werden Reize dargeboten (Kästchen), die jeweils drei Elemente (z.B. drei Zahlen) enthalten. Die Aufgabe besteht darin, jene Reize durchzustreichen, deren Elemente entweder eine aufsteigende oder eine abfallende Reihenfolge zueinander darstellen. Um Ablenkung zu erzeugen, wurden einzelne Itemelemente mit verschiedenen Zusatzmerkmalen versehen (z.B. Fettdruck, Hochstellung etc.), die für die Aufgabenbearbeitung irrelevant sind. Der TEMEKKO wurde in zwei Testformen (A und B) mit unterschiedlichen Inhaltsdomänen (figural, numerisch und verbal) entwickelt. Die Übertragung der Item-Elemente in die jeweilige Inhaltsdomäne erfolgte nach einer strikten Regel.

Die Auswertung des Verfahrens über die Verwendung des Konzentrationsleistungswertes (KL-Wert) ermöglicht eine nahezu verfälschungssichere Erfassung der individuellen Fähigkeit unter Berücksichtigung des Tempos und der Fehler gleichermaßen. Als zusätzliches Leistungsmaß kann der Fehler-Prozent-Wert ermittelt werden. Auswertungsschablonen sowie übersichtlich gestaltete Testformulare erhöhen die Ökonomie der Auswertung. Die gesamte Netto-Bearbeitungszeit beträgt 14 Minuten. Der TEMEKKO kann für Einzel- oder Gruppentestungen verwendet werden.

Die Validierung des TEMEKKO erfolgte in vier voneinander unabhängigen Studien. Drei der vier verschiedenen Stichproben umfassten überwiegend junge Erwachsene. Sie waren fast ausschließlich Studierende. Die vierte Stichprobe setzte sich aus Personen im fortgeschrittenen Alter (> 60 Jahre) zusammen.

Alle Daten wurden zunächst auf Plausibilität der Eingaben überprüft und sodann einer Analyse deskriptiver Statistiken unterzogen. Die Variablen wurden anschließend z-transformiert und Reaktionszeiten umgepolt. Nach Prüfung des Zusammenhangs mit Alter und Geschlecht wurde für deren Einflüsse kontrolliert, indem sie auspartialisiert wurden. Die Schätzung fehlender Werte erfolgte ggfs. über den EM-Algorithmus.

Die Prüfung der beiden Testformen A und B bzw. der TEMEKKO Untertests (figural, numerisch, verbal) auf deren Gleichheit zueinander wurde auf zweierlei Wegen vorgenommen, über Strukturgleichungsmodelle und alternativ über Kovarianzanalysen. Ein essentiell-paralleles Messmodell bildete beide Sachverhalte (Gleichheit der Testformen und Gleichheit der Materialarten) ab und wurde für Tempowerte wie für Fehler-Prozent-Werte gleichermaßen angewandt. In allen vier Studien zeigten sich akzeptable Modellfits sowohl für die Tempowerte als auch für die Fehler-Prozent-Werte. Es kann davon ausgegangen werden, dass die beiden Testformen A und B sowie die drei Materialarten (figural, numerisch, verbal) dasselbe Konstrukt jeweils gleich gut messen. Die Testformen A und B des TEMEKKO können als essentiell-parallele Formen anerkannt werden. Die alternative Prüfung anhand von Kovarianzanalysen wies lediglich aufgrund der Anwendung von überaus strengen Analysekriterien Abweichungen auf. Hinsichtlich der Prüfung der Testformen A und B auf deren Gleichheit zu einander wies der figurale Untertest in Studie 1 einen mittleren Effekt ($\eta^2 = .07$) auf. Dieser könnte durchaus aufgrund des fortgeschrittenen Alters der Probanden zustande gekommen sein, sodass bei

der Bearbeitung des allerersten Testblattes noch relativ niedrige Leistungen erbracht wurden. In Studie 4 zeigten die figuralen ($\eta^2 = .03$) und verbalen ($\eta^2 = .03$) Testformen Unterschiede, die allerdings als gering eingestuft wurden. Für die Prüfung der Materialarten auf Gleichheit zeigte sich ein geringfügig differenzierteres Bild. Der in Studie 1 gefundene Unterschied zwischen figuralem und verbalem Material war hinsichtlich seiner geringen Effektgröße ($\eta^2 = .001$) zu vernachlässigen. In Studie 2 waren die Effektstärken für die Paarungen numerisch : verbal ($\eta^2 = .02$) und figural : verbal ($\eta^2 = .02$) zwar gering, aber dennoch existent. Auch in Studie 3 wurden Unterschiede zwischen den Materialien festgestellt (figural : numerisch ($\eta^2 = .05$) und figural : verbal ($\eta^2 = .02$)) sowie in Studie 4 (figural : numerisch ($\eta^2 = .02$), numerisch : verbal ($\eta^2 = .02$) sowie figural : verbal ($\eta^2 = .08$)). Aufgrund dieser Effekte, die zum Teil im mittleren Bereich lagen, wurde der TEMEKKO als relativ materialunabhängig eingestuft. Diese Befunde veranlassten zur Bereitstellung unterschiedlicher Normen.

Die Durchführungsobjektivität kann bei einer präzisen Einhaltung der standardisierten Instruktion des TEMEKKO sowie durch Standardisierung der Untersuchungssituation gewährleistet werden. Die im TEMEKKO bindenden Auswertungsvorschriften stellen ein hohes Maß an Auswertungsobjektivität sicher. Eine Interpretationsobjektivität ist durch die bereitgestellten Normen gegeben.

Zur Schätzung der Reliabilitäten des TEMEKKO wurde die Parallelttestkorrelation für kleine Stichprobengrößen (nach Kristof, 1963) verwendet. Die Voraussetzungen zur Anwendung dieses Maßes waren insofern gegeben, als in modellhaften Analysen belegt war, dass die Testformen A und B essentiell-parallele Formen darstellten. Die Höhe der Reliabilitätsschätzungen für die Tempowerte des TEMEKKO Gesamtttests lagen über die vier Studien hinweg im Bereich von $r_{tt} = .93$ bis $r_{tt} = .97$ und wurden als sehr gut

eingestuft. Die numerischen ($r_{tt} = .89$ bis $r_{tt} = .96$) sowie die verbalen Untertests ($r_{tt} = .89$ bis $r_{tt} = .92$) wiesen ebenso sehr hohe Werte auf. Im figuralen Bereich lagen die Reliabilitäten im guten Bereich ($r_{tt} = .85$ bis $r_{tt} = .90$). Lediglich der Wert der Senioren-Stichprobe im figuralen Untertest ($r_{tt} = .79$) blieb unter den Erwartungen. Die Reliabilitätsschätzungen der Fehler-Prozent-Werte ergab über alle vier Studien und alle Inhaltsdomänen hinweg überraschend hohe Werte ($r_{tt} = .82$ bis $r_{tt} = .89$ für Auslassungsfehler-Prozente, $r_{tt} = .88$ für Fehler-Gesamt-Prozente). Die Reliabilitätsschätzungen der Verwechslungsfehler-Prozente lagen deutlich darunter ($r_{tt} = .51$ bis $r_{tt} = .89$), wenngleich sie auch unerwartet hoch waren. Dies ist ein Indiz für die Homogenität der Testitems sowie der verteilten Zusatzmerkmale und spricht für eine Form der Stabilität, mit der Fehler begangen werden.

Zur Schätzung der Konstruktvaliditäten des TEMEKKO wurden zunächst die Korrelationsmatrizen der Analysevariablen der einzelnen Studien in Augenschein genommen, um erste Hinweise auf Zusammenhänge der verwendeten Testvariablen zu erhalten. Mit Hilfe von Strukturgleichungsmodellen und linearen multiplen Regressionsanalysen bzw. latenten Regressionen wurden auf Konstruktebene wie auch auf funktionaler Ebene die Berechnungen zur Schätzung der Validität angestellt. Zur Verortung der TEMEKKO Testleistung wurden kognitive Leistungsmaße der Intelligenz, Aufmerksamkeit, Konzentration sowie der Arbeitsgedächtniskapazität verwendet.

In vier voneinander unabhängigen Studien wurde dargelegt, dass der TEMEKKO einen Merkmalsbereich erfasst, dem Konvergenz mit dem Konstrukt Konzentration unterstellt werden kann. Als Markertest des Faktors Executive Attention ($\lambda = .79$) (Studie 1) erfordert der Test vom Probanden Fähigkeiten, die eng verwoben sind mit der Auffassung von Konzentration, wie sie dem TEMEKKO zu Grunde liegt. Aus funktionaler Sicht lässt sich

die Fähigkeit, die Executive Attention darstellt, als Steuerung höherer kognitiver Fähigkeiten verstehen, i.e. als exekutive Kontrolle. Dies wurde durch die TEMEKKO Untertests als stärkste Indikatoren der exekutiven Kontrolle (Studie 2) deutlich ($\lambda = .89$ numerisch, $\lambda = .79$ verbal, $\lambda = .75$ figural). Ein gemeinsam geteilter Varianzanteil des TEMEKKO mit Konzentration auf latenter Ebene (74 % bzw. 81 %, Studie 3) sowie die rechnerische Unseparierbarkeit des TEMEKKO mit den Konzentrationstests d2 und FAKT (Studie 4) stützen die Annahme der Konstruktkonvergenz des TEMEKKO mit Konzentration.

Die Überprüfung der Zusammenhänge der analysierten Konstrukte zu Konzentration bzw. exekutiver Kontrolle anhand des TEMEKKO sowie untereinander verdeutlichte einmal mehr, wie eng diese ineinander verwoben, aber dennoch konzeptionell trennbar sind. Insbesondere waren hierbei die Nähe zwischen exekutiver Kontrolle und Relational Integration ($r = .80$) bzw. Konzentration (TEMEKKO) und Coordination ($r = .68$) hervorzuheben, sowie zwischen Konzentration (TEMEKKO) und Intelligenz ($r = .76$ bzw. $r = .77$). Wurde für allgemeine Intelligenz kontrolliert, so ließen sich die TEMEKKO Testleistung und Coordination in ähnlicher Höhe auf diesem Faktor verorten ($\lambda = .79$ bzw. $\lambda = .78$). Die Zusammenhänge der Konstrukte untereinander, wie zwischen Intelligenz und Coordination ($r = .74$) oder zwischen Intelligenz und Speichern im Kontext gleichzeitiger Verarbeitung ($r = .69$) untermauern die Nähe der Konstrukte.

Als weitere Belege für die Verortung der TEMEKKO Testleistung werden die Ergebnisse der Vorhersageanalysen gewertet. Als die wichtigsten Prädiktoren der TEMEKKO Testleistung über alle vier Studien hinweg galten insbesondere Konzentration und Intelligenz. Darüber hinaus kam der Umstellfähigkeit (shifting) eine nicht unerhebliche Bedeutung zu. Der TEMEKKO seinerseits galt als stärkster Prädiktor für intellektuelle

Leistungen und eignete sich deutlich besser zur Vorhersage von Intelligenz als andere Konzentrationstests.

Hinsichtlich der Inhaltsdomänen kann der TEMEKKO als relativ unabhängig von Materialeinflüssen bewertet werden. Lediglich der figurale Untertest zeigte zum Teil geringe, aber nicht zu vernachlässigende materialspezifische Anteile.

Normwerte wurden anhand von zwei von einander unabhängigen Stichproben (Schüler $N = 312$ und Senioren $N = 136$) ermittelt. Die Standardwerte für beide Stichproben liegen jeweils getrennt für die KL-Werte der Testformen A und B, für die jeweilige Inhaltsdomäne (figural, numerisch und verbal) sowie für die KL-Gestamwerte vor. Darüber hinaus wurden Normen für Fehler-Prozent-Werte der Testformen A und B sowie des Fehler-Prozent-Gesamtwertes bereitgestellt.

Inhaltsverzeichnis

1	Überblick.....	1
2	Theoretischer Hintergrund	3
2.1	Konzentration	3
2.2	Konzentrationstests.....	6
2.3	Typen von Konzentrationstests	7
2.3.1	Kennwerte der Konzentrationsleistung	10
2.3.2	Zusammenhänge verschiedener Konzentrationstests	12
2.4	Exekutive Funktionen.....	13
2.4.1	Exekutive Kontrolle.....	18
2.4.2	Typen von Tests zur Erfassung von exekutiven Funktionen.....	20
2.5	Zusammenhänge mit anderen kognitiven Konstrukten	21
2.5.1	Konzentration und Aufmerksamkeit	22
2.5.2	Konzentration und Mental Speed	27
2.5.3	Konzentration und Intelligenz	30
2.5.4	Konzentration und Arbeitsgedächtnis	31
3	Entwicklung des TEMEKKO	35
3.1	Itemkonstruktion.....	35
3.1.1	Items und Aufgabe	37
3.1.2	Fehlerarten	42
3.1.3	Kontrolle von instruktionswidrigem Bearbeiten	43
3.2	Gestaltung des Testformulars	45
3.3	Durchführung.....	46
3.4	Auswertung.....	49
4	Empirische Untersuchungen.....	52
4.1	Validierungsstudien	52
4.1.1	Auswertungsmethoden	52
4.1.2	Prozedere der Datenbearbeitung.....	53

5	Hauptgütekriterien	63
5.1	Objektivität	63
5.2	Reliabilität.....	63
5.3	Validität	66
6	Nebengütekriterien	70
6.1	Normierung.....	70
6.2	Ökonomie	72
6.3	Akzeptanz	72
7	Fazit	73
8	Perspektiven	74
9	Literaturverzeichnis.....	76
Anhang I:	Validierungsstudie 1.....	94
Anhang II:	Validierungsstudie 2	115
Anhang III:	Validierungsstudie 3	142
Anhang IV:	Validierungsstudie 4	176
Anhang V:	Itemverteilung	202
Anhang VI:	Normen	207

1 Überblick

Der **T**est zur **M**essung **e**xekutiver **K**ontrolle / **K**onzentration (TEMEKKO) wurde von Markus Bühner und Lothar Schmidt-Atzert entwickelt und im Zuge des Weiterentwicklungsprozesses aus theoretischen Überlegungen und diversen Vortests von Testteilen optimiert. Der TEMEKKO erfasst die Ausprägung der Konzentrationsleistung bzw. das Ausmaß an exekutiver Kontrolle bei Erwachsenen und kann als Einzel- oder Gruppentest durchgeführt werden.

Die Entwicklung von Konzentrationstests in der Vergangenheit zeigt, dass diese häufig unterschiedlichen psychologischen Fachdisziplinen entspringen (z.B. Arbeitspsychologie, klinische Psychologie, Pharmakopsychologie o.ä.), wobei zunächst stärker praktische als theoretische Fragestellungen im Vordergrund standen. Der TEMEKKO wurde theoriegeleitet entwickelt, validiert und normiert.

Bei der Konzeption des TEMEKKO wurde unter anderem versucht, sich an der DIN 33430 bzw. der DIN Screen zu orientieren. Die DIN 33430 regelt die „Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen“ (Kersting & Hornke, 2003). Sie wurde im Januar 2002 aufgrund der Initiative vom Berufsverband deutscher Psychologen (BdP) verabschiedet und ist ein Qualitätsmaßstab für den gesamten Prozess der Eignungsdiagnostik. Die Qualitätsanforderungen der DIN 33430 umfassen die an der Auswahl beteiligten Personen, ebenso wie die verwendeten Verfahren und enthaltenen Abläufe (Dries, Krumm & Lux, 2006). Die DIN Screen stellt einen Auszug aus der DIN 33430 dar und stellt einen Leitfaden zur Verfügung, mit dem ein Screening ermöglicht wird, ob ein konkreter Prozess oder Prozessabschnitt den Qualitätsstandards der

DIN 33430 genügt. Insbesondere relevant bei der Konstruktion des TEMEKKO war die Checkliste 1 der DIN Screen, in der die Anforderungen an Verfahrenshinweise zusammengefasst sind. Sie enthält 140 Aussagen (44 % aller Aussagen) der DIN 33430 spezifisch zur Konstruktion und empirischen Überprüfung sowie zur Anwendung, Auswertung und Interpretation eines Verfahrens (Kersting, 2006).

Der TEMEKKO kann aufgrund seiner theoriegeleiteten Konstruktion und seiner überzeugenden Reliabilitäten und Validitäten in der Forschung als auch in der Praxis gleichermaßen eingesetzt werden. Als praxisbezogene Einsatzgebiete kommen die Eignungsdiagnostik (Arbeits- und Organisationspsychologie) oder verkehrspsychologische Untersuchungen (z.B. für Berufskraftfahrer) in Betracht. Darüber hinaus ist es vorstellbar, den TEMEKKO in der pädagogischen Psychologie (Ein Drittel aller Kinder, die dem schulpsychologischen Dienst vorgestellt werden, haben eine vorläufige Diagnose „Konzentrationsprobleme“ (Berg & Westhoff, 2006).) oder in der klinischen Psychologie (Frontalhirnsyndrom oder -schädigung) einzusetzen, wenngleich für diese Personengruppen bislang keine Normwerte vorliegen. Ferner ist der TEMEKKO als Indikator für höhere kognitive Leistungen überall dort einsetzbar, wo herkömmliche Intelligenztests aufgrund ihrer Komplexität oder Länge nicht verwendet werden können, ohne den Anspruch zu erheben, Intelligenz zu erfassen.

2 Theoretischer Hintergrund

Zur Aufbereitung der theoretischen Grundlage des TEMEKKO seien zunächst die beiden Konstrukte *Konzentration* und *Exekutive Funktionen* dargestellt. Sie bilden die Kernbereiche des Erfassungsraumes des TEMEKKO, des Tests zur Messung exekutiver Kontrolle und Konzentration. Vertiefend wird jeweils auf Testtypen zur Erfassung dieser Fähigkeiten eingegangen. Zur Einbettung der beiden Konzepte unter anderen kognitiven Konstrukten werden in den darauf folgenden Abschnitten die theoretischen Schnittstellen und Gemeinsamkeiten mit den jeweiligen anderen kognitiven Konstrukten herausgearbeitet sowie mit empirischen Befunden gestützt.

2.1 Konzentration

Grundsätzlich kann Konzentration als Zustand oder als Persönlichkeitsmerkmal verstanden werden (Westhoff, 1995). Soll der Zustand Konzentration erfasst werden, so werden üblicherweise experimentalpsychologische Gruppenvergleiche unter Einbeziehung kontrollierter Umgebungsvariablen angestellt. Werden hingegen Störvariablen so umfassend wie möglich eliminiert und das Konstrukt über einen längeren Zeitraum hinweg erfasst, so handelt es sich um den Trait, das Persönlichkeitsmerkmal (Westhoff, 1995). Hinsichtlich dieser Kategorisierung konstatiert Bühner (2001) Konzentration als einen Zustand, dessen Messung mit situativen Faktoren (z.B. Ermüdung) und mit Personenmerkmalen (z.B. Alter, Geschlecht, Intelligenz o.ä.) kovariiert. Gleichwohl wird die Fähigkeit „sich zu konzentrieren“ als Persönlichkeitsmerkmal aufgefasst (Bühner, 2001, S. 19f.). Ungeachtet dessen sind Konzentrationsleistungen an das verwendete

Material (verbal, numerisch oder figural) sowie an die Aufgabenstellung gebunden (Schmidt-Atzert & Bühner 1997; Bühner, 2001). Aufgrund dieser Merkmale können Testleistungen entsprechend variieren.

Konzeptionell unterscheiden Schmidt-Atzert, Krumm und Bühner (2008) Konzentration nach den Bereichen *konzentrierte Aufmerksamkeit* und *konzentrierte Weiterverarbeitung*. Wie in Abschnitt 2.5.1 (Konzentration und Aufmerksamkeit) vertiefend dargestellt, repräsentiert konzentrierte Aufmerksamkeit den Überlappungsbereich der Konzepte Konzentration und Aufmerksamkeit. Der Wahrnehmungsbezug ist kennzeichnend für die Aufmerksamkeitsleistung, die Bearbeitung der Aufgaben unter erschwerenden Bedingungen für die konzentrierte Leistung. Bei der konzentrierten Weiterverarbeitung (z.B. Rechnen, Handlungsplanung etc.) steht nicht die Wahrnehmung im Vordergrund, sondern Informationsverarbeitungsprozesse, die durch entsprechend hohe Kontrollanforderungen charakterisiert sind (Goldhammer & Moosbrugger, 2006).

Zur Eingrenzung des Merkmalsbereiches Konzentration werden nachfolgend wiederkehrende Definitionsmerkmale zum Konstrukt aus der Literatur zusammengestellt und systematisiert.

Den Beginn einer konzentrativen Leistung stellt ein entsprechendes Handlungsziel (Zweck) dar. Ziel kann es sein, beispielsweise eine spezifische Reizkonfiguration zu entdecken und auf diese in einer bestimmten Weise (instruktionsgemäß) zu reagieren. Konzentration beginnt demnach mit der zweckgebundenen (intentionalen) Zuwendung zu Reizen (Beckmann, 1991; Beckmann & Strang, 1993; Berg, 1991; Berg & Imhof, 2001; Brickenkamp, 2002; Brickenkamp & Karl, 1986; Hoffmann, 1993; Neumann & Simon, 1994; Posner & Rafal, 1987; Westhoff, 1991, 1995). Ein Teil der Wahrnehmung wird

bewusst auf einen begrenzten Teil der Umgebung ausgerichtet (Selektion), wobei störende (für die Aufgabenbearbeitung irrelevante) Reize bewusst ausgeblendet bzw. abgeschirmt werden (Beckmann, 1991; Beckmann & Strang, 1993; Berg & Imhof, 2001; Brickenkamp, 2002; Brickenkamp & Karl, 1986; Hoffmann, 1993; Neumann & Simon, 1994). Die wahrgenommenen Reizkonfigurationen werden fortlaufend und zielorientiert beurteilt (Berg, 1991; Berg & Imhof, 2001; Brickenkamp, 2002; Brickenkamp & Karl, 1986; Westhoff, 1991) und entsprechende (instruktionsgemäße) Handlungsreaktionen werden mental koordiniert sowie meist motorisch ausgeführt (Berg, 1991; Berg & Imhof, 2001; Neumann & Simon, 1994; Posner & Rafal, 1987; Westhoff, 1991). Konzentriertes Arbeiten wird meist als anstrengend erlebt (Berg & Imhof, 2001; Posner & Rafal, 1987; Westhoff 1991, 1995) und von internen Zustandsvariablen (z.B. Ermüdung) oder externen Zustandsvariablen (z.B. Lärm) sowie von Persönlichkeitsvariablen (z.B. Leistungsmotivation oder Intelligenz) beeinflusst (Westhoff, 1995).

Die dem TEMEKKO zugrunde gelegte Begriffsbestimmung für den Merkmalsbereich Konzentration abstrahiert die zuvor genannten Charakteristika aus der Forschung und stellt somit eine top-down orientierte Testdefinition dar. Der TEMEKKO stützt sich auf folgende Definition: „Konzentration ist die Fähigkeit, unter Bedingungen, die das Erbringen einer kognitiven Leistung normalerweise erschweren, schnell und genau zu arbeiten.“ (Schmidt-Atzert, Büttner, Bühner, 2004; Schmidt-Atzert, Krumm, Bühner, 2008).

2.2 Konzentrationstests

Konzentrationstests lassen sich als objektive Leistungstests kategorisieren (Brähler, Holling, Leutner & Petermann, 2002). Sie sollen das jeweilige Konstrukt so rein wie möglich, unter Ausschluss beeinflussender Umgebungsvariablen (Störvariablen) erfassen (Lienert & Raatz, 1998). Dies ist aufgrund der Konzeption von Konzentration und anderer kognitiver Leistungen nicht uneingeschränkt möglich. In welchem Ausmaß auch immer, wird zusätzlich zur eigentlichen Konzentrationsleistung ein anderes Maß mit erfasst. Sei es eine Rechenleistung, Wahrnehmungs- oder Verarbeitungsgeschwindigkeit, motorische Fähigkeiten oder ein Anteil an schlussfolgerndem Denken, Gedächtnis etc. Dieses Task Impurity Problem (TIP) lässt sich nicht zur Gänze ausschließen. Gleichwohl gilt es die Datenanalyse möglichst auf latenter Ebene durchzuführen, um dieses Problem zu lindern. Bei der Konzeption der Testaufgaben ist darauf zu achten, dass Gedächtnisleistungen, Lern-, Transfer- und/oder Übungeffekte möglichst ausgeschlossen werden. Hinsichtlich näherer Details zu diesen Effekten sei auf die Arbeiten von Westhoff (1989), Westhoff und Dewald (1990), Westhoff und Hagemeyer (1992) oder Bühner (2001) verwiesen.

Ein Test zur Erfassung von Konzentration erfordert zusammenfassend (Bühner 2001) vom Probanden (1) die aktive, intentionale Zuwendung auf Reize (Westhoff 1995; Berg & Imhof 1998, vgl. auch Konzentration als Willensakt von Mierke 1957), (2) eine fortlaufende Reizselektion (Westhoff 1991, 1995), (3) die aufmerksame Wahrnehmung von Reizen (Westhoff 1991, 1995; Hoffmann 1993), (4) eine kontrollierte Verarbeitung von Reizen (Berg & Imhof, 1998), (5) die Abschirmung von Störungen bei der Bearbeitung einer Aufgabe (Berg & Imhof 1998, Neumann & Simon 1994), (6) die Beantwortung eines Reizes unter handlungsführenden Gesichtspunkten (Posner & Rafal 1987) sowie (7) Aufwendung energetischer Ressourcen (Berg & Imhof 1998). Diese

Merkmale stellen inhaltlich den Anforderungskatalog an die Aufgaben des TEMEKKO dar.

An dieser Stelle sei darauf hingewiesen, dass zwei Merkmalen aus der Literatur, die sich immer wieder zum Konstrukt Konzentration zugehörig finden (vgl. 2.1 Konzentration), lediglich unterstellt werden kann, in den Aufgaben des TEMEKKO operationalisiert zu sein. Es handelt sich hierbei um (1) die *bewusste* Ausblendung bzw. Abschirmung von irrelevanten Reizen (Beckmann, 1991; Beckmann & Strang, 1993; Berg & Imhof, 2001; Brickenkamp, 2002; Brickenkamp & Karl, 1986; Hoffmann, 1993; Neumann & Simon, 1994) sowie darum, dass (2) konzentriertes Arbeiten meist als anstrengend *erlebt* wird (Berg & Imhof, 2001; Posner & Rafal, 1987; Westhoff 1991, 1995). Eine Differenzierung in dieser Tiefe könnte im Zweifel nur durch Befragung der getesteten Person erhoben werden und wäre somit kaum objektivierbar.

2.3 Typen von Konzentrationstests

Tests zur Erfassung der Konzentration zählen, wie die Verfahren zur Messung von Aufmerksamkeit, Willensstärke sowie der allgemeinen psychischen Leistungsfähigkeit, zu den allgemeinen Leistungstests (Bartenwerfer, 1964, 1983). Grundsätzlich können drei Arten von Konzentrationstests unterschieden werden: (1) Durchstreichetests, wie z.B. der Test d2 (Brickenkamp 2002), (2) Rechentests, wie z.B. der Revisions-Test (Marschner, 1972) und (3) Sortierverfahren, wie z.B. der Konzentrations-Verlaufs-Test KVT (Abels, 1974). Durchstreichetests erfordern vom Probanden in erster Linie die Selektion von Reizen. Die Aufgabe von Durchstreichetests besteht darin, Reize mit vorgegebenen Zieleigenschaften (Targets) unter Reizen, die diese Eigenschaften nicht enthalten

(Distraktoren), zu entdecken und durchzustreichen. Mit Durchstreichetests wird eine kontinuierliche, zielbezogene und fokussierende Reizselektion erfasst (Bühner, 2001). Rechentests hingegen umfassen einfache kognitive Aufgaben, nämlich Rechnen. Auch Sortiertests verlangen einfache kognitive Operationen, wie das Beachten von Zahlen in einer Reihenfolge oder das Sortieren von Kärtchen. Der Konzentrations-Verlaufs-Test (KVT) (Abels, 1974; Bartenwerfer, 1964) ist das einzige standardisierte Sortierverfahren, das gegenwärtig im deutschsprachigen Raum verwendet wird (Bartenwerfer, 1964, 1983; Bühner, 2001). Ungeachtet seiner Bestimmung als Konzentrationstest wird der KVT von Sturm (2002) als Test zur Erfassung der visuellen Daueraufmerksamkeit (Aufmerksamkeitsintensität) eingestuft.

Über diese Einteilung der Konzentrationstests nach Durchstreiche-, Rechen- und Sortierverfahren hinaus, gibt es weitere Tests zur Erfassung der Konzentrationsleistung, die sich dieser Kategorisierung nicht zuordnen lassen. Es sind dies z.B. das Inventar Komplexer Aufmerksamkeit (INKA) von Heyde (2000), das Frankfurter Aufmerksamkeits-Inventar (FAIR) von Moosbrugger & Oehlschlägel (1996) oder der Zahlen-Symbol-Test aus dem Hamburg-Wechsler-Intelligenztest für Erwachsene (HAWIE) von Tewes (1991) sowie der Subtest 3 (Benennen von Farbwörtern unter Interferenzbedingungen) aus dem Farb-Wort-Interferenztest (FWIT) von Bäumler (1985). Darüber hinaus bezieht sich die Einteilung nach Durchstreiche-, Rechen- oder Sortierverfahren naturgemäß nur auf Paper-Pencil-Tests und nicht auf computerbasierte Verfahren.

Eine systematische Einordnung von Konzentrationstests wurde von Schmidt-Atzert, Bühner & Enders (2006) durchgeführt. Hierbei konnten der Zahlen-Symbol-Test aus dem Berliner-Intelligenz-Struktur-Test (BIS) (Jäger, Süß & Beauducel, 1997), der Zahlen-

Verbindungs-Test (ZVT) (Oswald & Roth, 1987), der Aufmerksamkeits-Belastungstest d2 (Brickenkamp, 2002) sowie der Frankfurter-Adaptive-Konzentrations-Test (FAKT) (Moosbrugger & Heyden, 1998) als Verfahren identifiziert werden, die geeignetermaßen Konzentration (als g-Faktor) erfassen. In dieser Untersuchung konnte auch gezeigt werden, dass das Konstrukt Konzentration in eine Facettenstruktur wie der des BIS-Modells der Intelligenz überführt werden kann.

Die Ausführungen dieses Abschnitts lassen zusammenfassend erkennen, warum der TEMEKKO als Durchstreichetest konzipiert wurde. Diese Form der Operationalisierung erlaubt es am besten, die Definitionsmerkmale des Konzepts Konzentration (vgl. 2.1 Konzentration) umzusetzen. Die Testaufgaben erfassen die Fähigkeit sich zu konzentrieren. Hierunter wird das Ausführen einer zielgerichteten fortlaufenden Selektion von Reizkonfigurationen unter Bedingungen verstanden, die normalerweise das Erbringen dieser Leistung erschweren. Die Erschwernis besteht darin, bei der Aufgabenbearbeitung irrelevante, störende Reize auszublenden. Bei der Konzeption als Rechentests ist es schwierig, Störreize mit aufzunehmen, die als erschwerende Bedingung (Ablenkung) bei der Bearbeitung der Aufgabe als Konzentrationstestaufgabe erforderlich ist. Damit kann eine wichtige Voraussetzung für Konzentration - der Selektionsaspekt der Wahrnehmung - nur schwer operationalisiert werden. Darüber hinaus könnte ein Konzentrationstest, der als Rechentest konzipiert ist, im Sinne einer Test(un)fairness möglicherweise Vorteile für Probanden beinhalten, die berufsmäßig häufig schnell kopfrechnen müssen (z.B. Kellner). Sortierverfahren (z.B. Karten) erlauben es wiederum nicht, eine Inhibition von vorschnellen Antworten zu erzeugen. Darüber hinaus müsste im Falle einer Gruppentestung gewährleistet sein, dass jeder Proband das gesamte Testmaterial zur Verfügung hat, was aus ökonomischen Gesichtspunkten untauglich wäre.

2.3.1 Kennwerte der Konzentrationsleistung

Als Leistungsmaß konzentrierten Arbeitens bei Durchstreiche- und Rechentests können sowohl die verarbeitete Reizmenge (Tempo-Werte) als auch die Anzahl der dabei gemachten Fehler (Fehler-Werte) angesehen werden (Brickenkamp, 2002; Brickenkamp & Karl, 1986; Debus, 1995; Westhoff, 1995). Die Anzahl der bearbeiteten Reize pro Zeiteinheit kann auch als Schnelligkeitsmaß aufgefasst werden (Bühner, 2001; Westhoff, 1995). Diese Anzahl bearbeiteter Items ist über die Intervalle hinweg unterschiedlich, was Brickenkamp (2002) als Schwankungsbreite bezeichnet und Westhoff (1995) als Variationsmaß der Leistung. Die Summe der Schnelligkeitsmaße über alle Zeitintervalle hinweg wird üblicherweise als Gesamtzahl-Wert (GZ-Wert) benannt (Brickenkamp, 2002; Bühner, 2001; Westhoff, 1995). Darüber hinaus ist es denkbar, die erreichte Gesamtzahl der bearbeiteten Zeichen pro Zeitintervall als Kurvenverlauf grafisch darzustellen (Westhoff, 1995).

Wenngleich die Bearbeitung der einzelnen Reize als einfach angesehen werden kann, treten dennoch aus verschiedenen Gründen Fehler auf. Es lässt sich hierbei zwischen Auslassungs- und Verwechslungsfehlern unterscheiden. Unter Auslassungsfehlern versteht man die Unterlassung der instruktionsgemäßen Markierung eines Zielobjekts. Bei Verwechslungsfehlern hingegen wird ein Distraktor fälschlicherweise als Target markiert (Brickenkamp, 2002; Schmidt-Atzert & Bühner, 1998). Werden Auslassungs- und Verwechslungsfehler über alle Zeiteinheiten hinweg aufaddiert, so spricht man vom Fehlerwert (Brickenkamp, 2002; Bühner, 2001; Westhoff, 1995). Um Fehlerwerte über verschiedene Individuen hinweg vergleichbar zu machen, werden sie am jeweils erreichten Gesamtzahl-Wert (GZ-Wert) durch Multiplikation mit 100 relativiert und als

Fehlerprozentanteil (F-Prozent-Wert) berichtet (Brickenkamp, 2002; Bühner, 2001; Westhoff, 1995).

Als Indikator für die Konzentrationstestleistung, der Tempo und Fehler gleichermaßen berücksichtigt, schlägt die Literatur zweierlei Maße vor. Zum einen die Gesamtzahl der bearbeiteten Zeichen abzüglich der Fehlerwerte (GZ - F Wert) und zum anderen die Anzahl richtig bearbeiteter Zielreize (Bühner, 2001; Westhoff, 1995). Letzterer wird als Konzentrationsleistungswert (KL-Wert) bezeichnet. Üblicherweise sind Gesamtzahl-Werte (GZ-Werte), Gesamtzahl-Werte abzüglich der Fehlerwerte (GZ - F Werte) und die Anzahl richtig bearbeiteter Zielreize (KL-Werte) hoch reliable Maße (Bühner, 2001), wohingegen Fehler- und Fehler-Prozent-Werte typischerweise nur über längere Zeiträume hinweg hohe Reliabilitäten aufweisen (Westhoff, 1995).

Bei Verwendung des KL-Wertes besteht der Vorteil, dass die Konzentrationstestleistung bei instruktionswidriger Bearbeitung weniger verfälscht wird, als dies bei der Anwendung des Gesamtzahl-Wertes abzüglich der Fehlerwerte (GZ - F Wert) der Fall ist (Bühner, 2001). Das absichtliche Überspringen (sog. "Ü-Syndrom", vgl. Brickenkamp, 2002, S. 23) von mehreren Aufgaben oder gar einigen Testzeilen, was durch einen hohen GZ-Wert ($Z > 115$) in Verbindung mit einem hohen Fehleranteil ($> 10\%$) bemerkt würde, führt folglich nicht zu einer besseren Testleistung. Anders, als bei dem hinlänglich bekannten Test d2 (Brickenkamp 2002), bei dem zur Ermittlung des KL-Wertes von der Anzahl der richtig markierten Targets die Verwechslungsfehler abgezogen werden, geschieht diese Subtraktion beim TEMEKKO nicht. Der Grund hierfür liegt darin, dass ein üblicherweise hoch reliables Maß (Anzahl richtiger Zielreize) nicht durch einen unreliablen Kennwert (Verwechslungen) verwässert werden soll.

Für das vorliegende Verfahren ist zusammenfassend festzuhalten, dass der Konzentrationsleistungswert (KL-Wert) und der Fehler-Prozent-Wert als maßgebliche Indikatoren der Konzentrationstestleistung angesehen werden (vgl. 3.4 Auswertung).

2.3.2 Zusammenhänge verschiedener Konzentrationstests

In der psychologischen Diagnostik werden verschiedene Verfahren als Konzentrationstests eingesetzt, denen jedoch kein einheitliches Konstrukt Konzentration zugrunde liegt (Bühner, 2001). Sie wurden in unterschiedlichen Bereichen der angewandten Psychologie entwickelt, wobei zunächst stärker praktische als theoretische Fragen im Vordergrund standen. Daher unterscheiden sich die Verfahren, die heute als Konzentrationstests verwendet werden, in ihren Aufgabenmaterialien (verbal, numerisch, figural) und -anforderungen (Durchstreiche-, Rechen- und Sortierverfahren) teilweise erheblich, sodass die Korrelationen der verschiedenen Konzentrationsleistungskennwerte stark variieren (Bartenwerfer, 1964, 1983; Bühner, 2001; Westhoff, 1995; Westhoff & Kluck, 1984). Die Zusammenhänge zwischen den Mengenleistungen (Tempowerte) verschiedener Konzentrationstests liegen im moderaten bis hohen Bereich; wohingegen sich jene zwischen den Fehler-Prozent-Werten im niedrigen bis moderaten Bereich bewegen (Bühner, 2001, S. 28). Die GZ-Werte der Rechentests korrelieren moderat bis hoch und jene der Durchstreichetests moderat. Als „allenfalls moderat“ bezeichnet Bühner (2001, S. 28) die Zusammenhänge der GZ-Werte von Durchstreiche- mit Rechentests.

Schmidt-Atzert und Bühner (1997) zeigten, dass mit zunehmender Komplexität der Aufgabenstellung und zunehmender Ähnlichkeit der dargebotenen Reize die Korrelationen zwischen den Tempomaßen verschiedener Konzentrationstests ansteigen. Die motorische

Fähigkeit, Zielobjekte rasch (instruktionsgemäß) durchzustreichen, war dabei nur unwesentlich an der Konzentrationstestleistung im Test d2 und im Fahnentest (unveröffentlichte Eigenentwicklung von Schmidt-Atzert & Bühner, 1997) beteiligt. Hingegen war die Fähigkeit, einfache Reize zu diskriminieren (operationalisiert durch d und p), maßgeblich an der Testleistung im Test d2 und im Fahnentest beteiligt.

Moderate oder gar hohe empirische Zusammenhänge zwischen Konzentrationstestkennwerten stellen keine Besonderheit bei Konzentrationstests dar. Auch Intelligenz-Subtests etwa korrelieren mit Konzentrationstestkennwerten im moderaten bis hohen Bereich (Bühner, 2001; Westhoff & Kluck, 1983).

2.4 Exekutive Funktionen

Exekutive Funktionen stellen auf funktionaler Ebene ein System zur Überwachung und Regulation der Informationsverarbeitung und des Verhaltens dar. Exekutive Funktionen subsumieren einzelne Aktivitäten, die Individuen dazu befähigen, planmäßig aber flexibel, zielgerichtet und effektiv zu handeln. Sie beschreiben eine Vielzahl verschiedener Kontrollprozesse, die gleichzeitig miteinander in Verbindung stehen und sich wechselseitig beeinflussen (Miyake, Friedman, Emerson, Witzki & Howerter, 2000; Monsell & Driver, 2000). Zu den exekutiven Funktionen zählen nach Goschke (2002) das Arbeitsgedächtnis, das eine aktive Aufrechterhaltung aufgabenrelevanter Informationen ermöglicht, sowie die Inhibition automatisierter Antworten bzw. inadäquater Reaktionen. Darüber hinaus wird die Steuerung von Aufmerksamkeitsprozessen hinzu gezählt, insbesondere die Fähigkeit, die Aufmerksamkeit auf bestimmte Informationen zu fokussieren (fokussierte Aufmerksamkeit), sowie der Wechsel der Aufmerksamkeit von einer Bezugsquelle zu

einer anderen (Aufmerksamkeitsflexibilität) (Roberts, 2003; Monsell & Dirver, 2000; Schneider, Owen & Duncan, 2000). Neben der Aufgabe, Aufmerksamkeitsprozesse zu steuern, organisiert das exekutive System Lernprozesse und Strategien zur Problemlösung und ist damit sowohl bei abstrakten Denkprozessen der Entscheidungsfindung sowie im Dienste der vorausschauenden Handlungsplanung aktiv als auch bei der Ausführung, deren zeitlicher Strukturierung, der Fehlererkennung und Fehlerkorrektur (Baddeley & Della Salla, 2003; Carlson, 2003; Roberts, 2003; Andreasen, 2002.) Nachfolgend seien jene drei exekutiven Funktionen herausgestellt, die meist in Forschungsarbeiten eingehen; i.e. *Shifting*, *Updating* und *Inhibition*.

Unter *Shifting* wird eine Fähigkeit beschrieben, die den Wechsel zwischen mehreren Aufgaben, Operationen oder Einstellungen (mental sets) darstellt (Monsell, 1996). Sie wird in der Literatur oft auch als Aufmerksamkeitswechsel (attentional switching) oder Aufgabenwechsel (task switching) bezeichnet. Darüber hinaus wird in Modellen der Aufmerksamkeitskontrolle (attentional control), wie z.B. dem Supervisory Attentional System Modell (SAS) (Norman & Shallice, 1986), angenommen, dass die Fähigkeit zwischen Aufgaben oder Einstellungen zu wechseln, einen wichtigen Aspekt der exekutiven Kontrolle darstellt (Miyake, Friedman, Emerson, Witzki & Howerter, 2000). Befunde aus früheren Studien zeigen, dass der Wechsel zwischen Einstellungen (mental sets) in zusätzlichem Zeitaufwand (i.S.v. verlängerten Reaktionszeiten) resultiert (z.B. Jersild, 1927; Rogers & Monsell, 1995), insbesondere, wenn der Wechsel internal angestoßen werden muss, als durch externe Reize (Spector & Biederman, 1976). Eine naheliegende, wenn auch unzutreffende Annahme wäre, dass das Disengagement von einer Aufgabe und das entsprechende Engagement zu einer neuen Aufgabe die Kosten (i.S. verlängerter Reaktionszeiten) verursachen könnten. Vielmehr gilt es, proaktive Interferenz oder negatives Priming der vorauslaufenden Aufgabe beim selben Stimulustyp zu

überwinden (Wylie & Allport, 2000). Diese Form von Wechsel (shifting) ist nicht bedeutungsgleich mit der Fähigkeit der räumlichen Verlagerung oder dem Wechsel von visueller Aufmerksamkeit bei bewussten Augenbewegungen oder verdeckter visueller Aufmerksamkeit (Posner & Raichle, 1994). Der Wisconsin Card Sorting Test (Grant & Berg, 1981) erwies sich bei Miyake et al. (2000) als am engsten mit Shifting verknüpft. Typischerweise werden Shifting-Funktionen bzw. kognitive Flexibilität durch Switching Aufgaben erfasst.

Updating (Updating and Monitoring of Working Memory Representations) wird in seiner Bedeutung eng mit dem Arbeitsgedächtnis in Verbindung gebracht (Jonides & Smith, 1997; Lehto, 1996). Updating erfordert die Überwachung und Kodierung eingehender Informationen hinsichtlich ihrer Relevanz für die zu bewältigende Aufgabe, um dann die Items, die im Arbeitsgedächtnis gehalten werden, entsprechend zu überprüfen, i.e. alte, nicht mehr relevante Items werden durch neue, relevante ersetzt (Morris & Jones, 1990). Updating geht über simple Aufrechterhaltung von Information hinaus, da es die aktive Manipulation relevanter Arbeitsgedächtnisinhalte erfordert (Lehto, 1996; Morris & Jones, 1990). Operation-Span-Aufgaben werden typischerweise zur Erfassung von Updating verwendet.

Inhibition (Inhibition of Prepotent Responses) beschreibt die Fähigkeit, bewusst dominante, automatische oder vorherrschende Antworten zu unterdrücken, wenn erforderlich. Reaktive Inhibition, wie negatives Priming oder Inhibition of Return sind davon abzugrenzen, da sie üblicherweise nicht beabsichtigte Prozesse darstellen (Logan, 1994). Das Verfahren, das typischerweise Inhibition erfasst, ist der Stroop Test (Stroop, 1935) (bzw. der Farb-Wort-Interferenztest (FWIT) von Bäumlner (1985)); gleichwohl wird dieser auch zur Erfassung der Konzentrationsfähigkeit eingesetzt.

Miyake, Friedman, Emerson, Witzki und Howerter (2000) zeigten, dass die exekutiven Funktionen Shifting, Updating und Inhibition zwar moderat korrelierten ($r = .42$ bis $r = .63$), nichtsdestotrotz aber klar voneinander getrennt werden konnten. Auch bei Friedman, Miyake, Corley, Young, De Fries & Hewitt (2006) ließen sich Zusammenhänge im Bereich von $r = .39$ bis $r = .64$ zeigen, wobei diese aufgrund der Berechnungen mit Strukturgleichungsmodellen doppelt minderungskorrigiert, d.h. stark aufgewertet, waren.

In diesem Zusammenhang sei darauf hingewiesen, dass exekutive Funktionen sich naturgemäß erst dann offenbaren, wenn sie andere kognitive Funktionen steuern. Jede exekutive Aufgabe beinhaltet andere für die exekutive Funktion selbst nicht zwangsläufig maßgebliche kognitive Prozesse. Miyake, Friedman, Emerson, Witzki & Howerter (2000) weisen darauf hin, dass korrelative Zusammenhänge zwischen exekutiven Funktionen mit Vorsicht zu interpretieren sind. Untersuchungen individueller Unterschiede über verschiedene gesunde und klinische Stichproben hinweg, in denen Verfahren wie der WCST Wisconsin Card Sorting Test (Grant & Berg, 1964) oder der TOH Tower of Hanoi Test (Simon, 1975) verwendet wurden, zeigen durchgängig Interkorrelationen von $r = .40$ oder weniger, die oft nicht signifikant werden. Solche Untersuchungen, die üblicherweise einen korrelativen oder faktorenanalytischen Ansatz verwenden, weisen deutliche Schwächen auf (Baddeley, Della Salla, Gray, Papagano & Spinnler, 1997; Rabbit, 1997a). Demnach müssen niedrige Korrelationen zwischen exekutiven Funktionen oder mehrere separierbare Faktoren nicht unbedingt auf tatsächlich trennbaren exekutiven Funktionen beruhen (Miyake & Shah, 1999). Möglicherweise maskieren nicht-exekutive Anforderungen der Verarbeitung (z.B. Sprache oder visuell-räumliche Verarbeitung) zu Grunde liegende Gemeinsamkeiten der exekutiven Aufgaben. Ebenso kann ein niedriger Leistungswert bei einer einzelnen exekutiven Aufgabe nicht unbedingt als Ineffizienz oder

Beeinträchtigung des exekutiven Systems interpretiert werden. Das hier angesprochene Task Impurity Problem ist methodisch zu kontrollieren, mindestens aber bei der Interpretation von Analyseergebnissen zu berücksichtigen.

Friedman, Miyake, Corley, Young, De Fries und Hewitt (2006) berichten über Zusammenhänge zwischen exekutiven Funktionen und Intelligenz. Überraschenderweise fanden sie den höchsten Zusammenhang mit fluider Intelligenz bei Updating, nicht aber bei Shifting und Inhibition. Auch kristalline Intelligenz zeigte deutliche Zusammenhänge mit allen drei exekutiven Funktionen, den höchsten mit Updating. Selbst wenn Inhibition und Shifting in den Modellen unberücksichtigt blieben, verschlechterte dies das Modell im Vergleich zu kontrollierten Interkorrelationen der exekutiven Funktionen nicht. Man kann daraus schließen, dass exekutive Funktionen in unterschiedlichem Ausmaß mit Maßen der Intelligenz zusammenhängen. Weitere Analysen ergaben, dass insbesondere Updating substantiell zum Zusammenhang zwischen fluider und kristalliner Intelligenz beitrug. Nicht unerwähnt soll bleiben, dass ein erheblicher Varianzanteil der Intelligenzmaße unaufgeklärt blieben (49 % bei kristalliner Intelligenz, 57 % bei fluider Intelligenz). Folglich sind exekutive Funktionen – in unterschiedlichen Gewichtungen - wichtig in der Erbringung intellektueller Leistungen, aber nicht allein maßgeblich.

Der starke Zusammenhang von Intelligenz und Updating deckt sich mit Befunden über Zusammenhänge zwischen Intelligenz und der Arbeitsgedächtniskapazität (Carpenter, Just & Shell, 1990; Engle, Tuholski, Laughlin & Conway, 1999). Updating und Arbeitsgedächtniskapazität werden als eine Fähigkeit beschrieben, die attentionale Kontrolle erfordern, um relevante Informationen und Aufgabenziele aufrecht zu erhalten, trotz Interferenzen diese Informationen zu löschen, sollten sie irrelevant werden, und sie durch neue Informationen zu ersetzen (Engle et al., 1999; Miyake, Friedman, Emerson,

Witzki & Howerter, 2000). Diese Ausführungen überlappen begrifflich stark mit der Auffassung von Intelligenz nach Binet (zit. nach Carroll, 1993; p. 35). Demnach besteht Intelligenz aus zwei Hauptprozessen; zunächst die Wahrnehmung der externalen Welt, danach das Wiedereinsetzen der Wahrnehmungen in das Gedächtnis, die Überarbeitung und das Nachdenken darüber.

Die Beschreibungen der exekutiven Funktionen Shifting und Inhibition lassen deutliche Übereinstimmungen mit der Aufmerksamkeitskomponente Aufmerksamkeitswechsel bzw. Konzentration erkennen. Sturm (2002) versteht unter dem Aufmerksamkeitswechsel oder kognitiver Flexibilität den schnellen Wechsel des Aufmerksamkeitsfokus zwischen verschiedenen Informationsquellen, also ebenso eine Shifting-Fähigkeit.

2.4.1 Exekutive Kontrolle

Die Fähigkeit, die einzelne exekutive Funktionen steuert und reguliert, ist die Exekutive Kontrolle. Die frühe Forschung zu exekutiver Kontrolle basierte auf der Unterscheidung zwischen automatischen und kontrollierten (effortful) Prozessen oder gewohnheitsmäßigen und neuen Aktivitäten (Shiffrin & Schneider, 1977). Automatische Prozesse erfordern wenig attentionale Ressourcen, kontrollierte Prozesse dahingegen mitunter viel. Exekutive Kontrolle reguliert die Informationsverarbeitung und Reaktionsauswahl (response selection) in Situationen, in denen automatisierte Mechanismen entweder nicht vorhanden sind, oder für die Aufgabenerfüllung nicht adäquat sind (Norman & Shallice, 1986). Solche Situationen sind z.B. Entscheidungsfindung, Unterdrückung von gewohnheitsmäßigem Antwortverhalten, Fehlerkorrekturen oder generell neue und schwierige Situationen. Norman & Shallice (1986) postulieren, dass das Supervisory

Attentional System (SAS) eine Steuerungsstelle für die Auswahl geeigneter Schemata in solchen Situationen darstellt. In der Literatur werden Konzepte, wie Frontallappen-Funktionen, kognitive Kontrolle oder attentionale Kontrolle nahezu synonym mit exekutiver Kontrolle verwendet. Kognitive Kontrolle wird als Fähigkeit verstanden, Handlungen und Gedanken in Übereinstimmung mit internalen Absichten anzuleiten (Cohen, Botvinick & Cartner, 2000). Darüber hinaus wird dem Konzept eine begrenzte Kapazität zugeschrieben und ermöglicht eine top-down Verarbeitung von aufgabenrelevanten Reizen (Miller und Cohen, 2001).

Logan (2003) beschreibt exekutive Kontrolle als einen paradoxen Zustand, nämlich als einen Prozess, bei dem der Verstand (mind) sich selbst kontrolliert. Der Zweck der exekutiven Kontrolle besteht im Erreichen einer Kohärenz und optimalen Organisation der kognitiven Aktivität und deren Umsetzung in zielgerichtetes Verhalten. Exekutive Kontrolle ist erforderlich, wenn das kognitive System von einer Anforderung auf eine andere umgestellt werden muss (shifting), wenn Prozesse der Wahrnehmung oder der Handlungskontrolle vorzubereiten sind (planning), wenn habituelle Verhaltensantworten auf Stimuli unterdrückt werden müssen (inhibition), wenn bereits eingeleitete Verhaltensantworten modifiziert oder abgebrochen werden müssen (updating), oder wenn kognitive Operationen ausgewählt (updating) oder koordiniert werden müssen. Die Beteiligung der exekutiven Kontrolle am exekutiven System wird als am stärksten erachtet, wenn die Aufgabe neu ist (Rabbitt, 1997b).

Das Aufgabenwechselparadigma wird typischerweise zur Untersuchung exekutiver Kontrollprozesse herangezogen. Hierbei muss zwischen unterschiedlichen Aufgaben rasch gewechselt werden. Der Annahme, dass die Umstellfähigkeit eine exekutive Funktion erfordert, liegt der Gedanke zugrunde, dass bei einem Wechsel eine neue

Aufgabenkonfiguration (task set) aktiviert wird, während eine bestehende, konkurrierende Konfiguration unterdrückt werden muss (Monsell, 1996). Die Rolle der exekutiven Prozesse liegt hierin, dass Aktivierungs- und Hemmprozesse initiiert und ein Gleichgewicht zwischen beiden Prozessen hergestellt werden muss, um eine optimale Leistung zu erbringen. Logan (2004) bezeichnet die Fähigkeit zum flexiblen Wechseln als wesentlichstes Kennzeichen von exekutiver Kontrolle.

2.4.2 Typen von Tests zur Erfassung von exekutiven Funktionen

So vielfältig verschiedene mentale Aktivitäten eines Individuums sind, die unter exekutiven Funktionen verstanden werden können (vgl. 2.4 Exekutive Funktionen), so facettenreich können die Operationalisierungen und Paradigmen derselben ausfallen. Darüber hinaus sind die Verfahren entsprechend ihrer konzeptionellen Zuordnung mitunter als Tests zur Erfassung der Arbeitsgedächtniskapazität oder der Aufmerksamkeit oder der Intelligenz gelistet. Nichtsdestotrotz können einige Tests als Verfahren benannt werden, die typischerweise zur Erfassung von exekutiven Funktionen eingesetzt werden.

So kommt der Tower of Hanoi (TOH) (Simon, 1975) als neuropsychologischer Test zur Erfassung von Planungs- und Problemlösefähigkeiten zum Einsatz. Zur Messung der kognitiven Flexibilität (shifting) wird häufig der Wisconsin Card Sorting Test (WCST) (Grant & Berg, 1981) verwendet. Beiden Verfahren wird allerdings eine kaum fundierte Konstruktvalidität vorgeworfen (Phillips, 1997; Rabbitt, 1997b). Ebenso operationalisiert der Subtest Switching Numerisch aus der Arbeitsgedächtnistest-Batterie von Oberauer, Süß, Schulze, Wilhelm und Wittmann (2000) bzw. die Switching-Kosten-Aufgaben der Arbeitsgedächtnistest-Batterie von Oberauer, Süß, Wilhelm und Wittmann (2003) das

Switching-Paradigma. Operation Span Aufgaben, wie z.B. die Letter-Memory-Task (Morris & Jones, 1990) oder Reading Span (Oberauer et al., 2000) sind funktionell Updating zuzuschreiben. Zur Erfassung von Inhibition wird klassischerweise der Stroop-Test (Stroop, 1935) verwendet, aber auch andere Verfahren, wie z.B. die Stop-Signal-Aufgabe (Logan, 1994) oder die Antisakkaden-Aufgaben (Roberts, Hager, & Heron, 1994). Miyake, Friedman, Emerson, Witzki & Howerter (2000) konnten der Inhibition den Tower of Hanoi Test (TOH) zuordnen.

Ein Verfahren, das konzeptionell zur Erfassung exekutiver Kontrolle geeignet ist, misst Fähigkeiten, die für die Steuerung der Informationsverarbeitung sowie für die Auswahl der geeigneten Beantwortung eines Items erforderlich sind. Darüber hinaus wird jede Aufgabe als neuartiger Reiz konzipiert, sodass nicht auf automatische Prozesse bei der Aufgabenbearbeitung zurückgegriffen werden kann. Ferner werden die Aufgaben mittels Störreizen so gestaltet, dass sie Ablenkung erzeugen und vorschnelles Antworten evozieren. Erst die Kontrolle und ggfs. Unterdrückung der Antworttendenz ermöglicht, dass das jeweilige Item richtig beantwortet werden kann. Diese Forderungen werden für den TEMEKKO postuliert.

2.5 Zusammenhänge mit anderen kognitiven Konstrukten

In den nachfolgenden Abschnitten werden die Zusammenhänge, Überlappungen und Abgrenzungen zu Konstrukten, wie Aufmerksamkeit, Mental Speed und Intelligenz (nach dem Modell des Berliner-Intelligenzstruktur-Tests (BIS) (Jäger, Süß & Beauducel, 1997), sowie zur Arbeitsgedächtniskapazität konzeptionell und empirisch dargestellt.

2.5.1 Konzentration und Aufmerksamkeit

In der Alltagssprache werden die Begriffe *Konzentration* und *Aufmerksamkeit* häufig synonym verwendet (Bühner, 2001; Westhoff, 1995). Auch in der Wissenschaft werden die beiden Konzepte oft nicht klar voneinander abgegrenzt (Berg & Imhof, 2001; Bühner, 2001). Semantisch können sie insofern unterschieden werden, als dass sich Aufmerksamkeit auf die Wahrnehmung und Konzentration auf das Arbeiten bezieht (Freyberg, 1989, zitiert nach Westhoff, 1995).

Im Allgemeinen versteht man unter Aufmerksamkeit die Fähigkeit des Wahrnehmungssystems, bestimmte Reize aus einer Flut einströmender Umgebungsreize zu selektieren und dabei irrelevante (störende) Reize auszublenden (Berg & Imhof, 2001, Schmidt-Atzert, Büttner & Bühner, 2004, Westhoff, 1995). Im Modell der Aufmerksamkeit von van Zomeren und Brouwer (1994) werden die Aspekte Intensität und Selektivität unterschieden. Die Selektion kann entweder intentional (goal-directed) oder nicht-intentional (stimulus-driven) erfolgen und kann zu einer Reaktion führen. In der Literatur werden verschiedene Arten der Aufmerksamkeit unterschieden, i.e. selektive, fokussierende, gerichtete und geteilte Aufmerksamkeit nach dem Selektivitätsaspekt sowie Daueraufmerksamkeit und Vigilanz nach dem Intensitätsaspekt. Diese unterschiedlichen Benennungen werden zum Teil synonym verwendet (Bühner, 2001; Schmidt-Atzert, Krumm & Bühner, 2008; Westhoff, 1995). Als fokussierte Aufmerksamkeit benennt Sturm (2002) jene „Fähigkeit, einen spezifischen Realitätsausschnitt zu isolieren, um ihn einer differenzierteren Analyse zu unterziehen. Dabei ist es erforderlich, den Fokus auch unter ablenkenden Bedingungen aufrecht zu erhalten und die Interferenz durch parallel ablaufende, automatische Verarbeitungsprozesse zu unterdrücken“ (Sturm, 2002, S. 108) An dieser Stelle wird bereits die Parallele zum Konstrukt der Konzentration deutlich.

Diese Einengung des Wahrnehmungsraumes kann wiederum danach unterschieden werden, ob sie gerichtet oder ungerichtet ist. Wird die Wahrnehmung nicht mit der Absicht nach einem spezifischen Ausschnitt des Wahrnehmungsraumes eingegrenzt, sondern aufgrund bestimmter Informationen (z.B. spezifische Eigenschaften eines Zielreizes), so spricht man von selektiver Aufmerksamkeit. Wird darüber hinaus die Wahrnehmung auf einen bestimmten Realitätsausschnitt eingeschränkt und zusätzlich auf bestimmte Informationen aus einer anderen Sinnesmodalität geachtet (akustisch), so spricht man von geteilter Aufmerksamkeit. Als Daueraufmerksamkeit oder Vigilanz betrachtet man jene Einengung der Wahrnehmung auf einen Teil der Umgebung, wenn sie über einen längeren Zeitraum (z.B. mehrere Stunden) erfolgt und die Aufgabe darin besteht, auf selten auftretende Zielreize adäquat (instruktionsgemäß) zu reagieren (Schmidt-Atzert, Krumm & Bühner, 2008; Westhoff, 1995). Daueraufmerksamkeit wird üblicherweise mit lange andauernden Signalentdeckungsaufgaben erfasst, in denen ein hoher Anteil an relevanten Stimuli enthalten ist, wohingegen Aufgaben zur Erfassung der Vigilanz ebenfalls lange andauernd und monoton sind, aber einen niedrigen Anteil an relevanten Stimuli haben (Sturm, 2004). Vigilanzbedingungen stellen daher höhere Anforderungen an die kognitive, internale Kontrolle. Zur Vertiefung auf bedeutende Strukturmodelle der Aufmerksamkeit sei auf Schmidt-Atzert, Krumm & Bühner (2008) verwiesen.

Die entscheidende Schnittstelle ist der Selektivitätsaspekt beider Konstrukte. Selektive Aufmerksamkeit konnte als stabiler Faktor (Reaktionszeitmediane der TAP Untertests Go/Nogo und Geteilte Aufmerksamkeit getrennt für Quadrate und Töne) wiederholt über verschiedene Forschungsarbeiten hinweg repliziert werden (Bühner, Schmidt-Atzert, Grieshaber & Lux, 2001; Bühner & Schmidt-Atzert, 2002; Zimmermann & Fimm, 2002) und stellt das Kernkonzept der Aufmerksamkeit dar. Konzentration und Aufmerksamkeit beinhalten gleichermaßen, dass beim Probanden eine Reizselektion unter Abschirmung

irrelevanter Reize stattfindet (Berg & Imhof, 2001; Bühner, 2001; Häcker & Stapf, 1998; Moosbrugger & Oehlschlägel, 1996; Neumann & Simon, 1994; Rapp, 1982; Westhoff, 1995). Bei einer konzentrativen Leistung erfolgt die Selektion notwendigerweise willentlich (goal-directed), wohingegen diese bei Aufmerksamkeitsprozessen auch unwillkürlich (stimulus-driven) verlaufen kann. Dessen ungeachtet erfordert die Wahrnehmung oder Selektion von Reizen ebenso ein gewisses Maß an Konzentration, zumindest bei komplexen oder wenig automatisierten Aufgaben. Die Differenzierung beider Konstrukte stützen Schmidt-Atzert, Krumm, Bühner (2008) auf die Annahme, Aufmerksamkeit auf den Wahrnehmungsprozess zu beschränken, wohingegen Konzentration den gesamten Informationsverarbeitungsprozess umfasst, von der Wahrnehmung eines Reizes bis einschließlich der geeigneten Reaktionsauswahl (response selection). Darüber hinaus läuft bei der Konzentration die Reizselektion und Weiterverarbeitung kontinuierlich ab, was dem Individuum Anstrengung abverlangt.

Hieraus lassen sich zweierlei Haltungen konstatieren. Zum einen wird Konzentration zuweilen als Spezialfall der Aufmerksamkeit gesehen. Dann nämlich, wenn Aufmerksamkeit auf aufgabenrelevante Reize willentlich gerichtet wird (Moosbrugger & Oehlschlägel, 1996; Neumann & Simon, 1994; Sturm 1997). Zum anderen stellt Aufmerksamkeit zwar einen Bestandteil bei der Erbringung konzentrativer Leistungen dar, wird aber als ein davon unabhängiger Prozess gesehen, da Konzentration über den Wahrnehmungsprozess hinausgeht (Berg & Imhof, 2001; Schmidt-Atzert, Krumm, Bühner, 2008).

Zur Verdeutlichung der konzeptionellen Abgrenzung bzw. Überlappung von Konzentration und Aufmerksamkeit sei nachfolgend (Abbildung 1) das Modell nach Schmidt-Atzert, Krumm & Bühner (2008) dargestellt. Es stellt einen Minimalkonsens der

gängigen rationalen und (induktiv) empirischen Ansätze dar. Das Modell erhebt nicht den Anspruch, die Konstrukte Aufmerksamkeit und Konzentration in ihren Funktionen vollumfänglich abzubilden, sondern stellt den Status des aktuell Messbaren heraus. *Konzentrierte Aufmerksamkeit* stellt den Überschneidungsbereich beider Konzepte dar. Aufmerksamkeitsleistungen unter erschwerten Bedingungen, wie z.B. Zeitdruck, Störreize, Lärm, lange Dauer etc., kennzeichnen diesen Aspekt. Testaufgaben erfordern schnelles Durchsuchen nach Targets (Fokus- bzw. Selektionsaspekt) und schnelles Antworten unter bestimmten Testbedingungen, wie die gleichzeitige Darbietung aller Reize und sukzessive, schnelle Bearbeitung (Execute-Aspekt). Es sind Aufgaben, die die Aufmerksamkeitsanforderungen erfüllen, wie auch willentliche Anstrengung abverlangen (Goldhammer & Moosbrugger, 2006; Schmidt-Atzert, Büttner & Bühner, 2004). Dessen ungeachtet, kann die Komponente konzentrierte Aufmerksamkeit auch Anforderungen an die Komponenten Alertness, Selektion oder Daueraufmerksamkeit stellen, nämlich dann, wenn diese unter willentlicher Anstrengung (cognitive effort) zu erbringen sind. Die Komponente konzentrierte Weiterverarbeitung ist konzeptionell von der Aufmerksamkeit abzugrenzen. Sie beinhaltet Informationsverarbeitungsprozesse mit hohen Kontrollanforderungen.

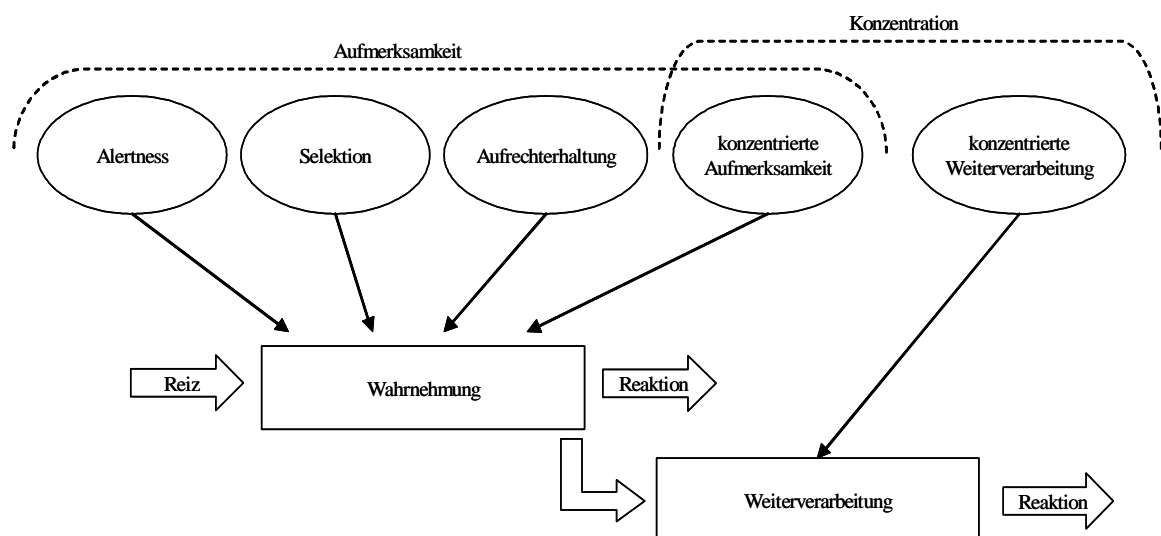


Abbildung 1: Konzentrationsmodell von Schmidt-Atzert, Krumm und Bühner (2008)

Tests zur Erfassung der Aufmerksamkeit werden häufig computergestützt konzipiert. Die Reize werden meist in einer bestimmten festgelegten Geschwindigkeit (force paced) auditiv und/oder visuell dargeboten. Die Beantwortung der in der Instruktion spezifizierten Targets erfolgt verbal oder manuell. Bei Verfahren zur Erfassung der geteilten Aufmerksamkeit werden Versuchspersonen simultan in zwei Sinnesmodalitäten (akustisch und optisch) Reizen ausgesetzt (Bühner, 2001; Westhoff, 1995). Die Aufgaben in Vigilanztests sind entweder sukzessive oder simultane Diskriminationsaufgaben. Bei sukzessiven Aufgaben werden nacheinander gleichförmige Stimuli dargeboten, unter denen sich – wenn auch mitunter selten – ein ungleichförmiger Reiz befindet, auf den instruktionsgemäß zu reagieren ist. Simultane Diskriminationsaufgaben stellen die Anforderung, dass Probanden unter mehreren gleichzeitig dargebotenen Stimuli einen Zielreiz entdecken müssen (Bühner, 2001). Die Verfahren erfassen die benötigten Reaktionszeiten und Fehlerraten der Testpersonen (Bühner, 2001; Westhoff, 1995). Im Gegensatz zu der meist festgelegten Taktung der Stimuluspräsentation bei Aufmerksamkeitstests können bei Konzentrationstests die Probanden selbst die Bearbeitungszeit pro Item (innerhalb einer vorgegebenen Gesamtbearbeitungszeit für das gesamte Verfahren oder Testteile) festlegen (self paced). Paper-Pencil-Tests der Aufmerksamkeit bestehen meist aus einfachen Suchaufgaben, bei denen Targets unter Distraktoren entdeckt werden müssen und instruktionsgemäß darauf zu reagieren ist. Auch hier bestimmen die Probanden die Bearbeitungsgeschwindigkeit der einzelnen Reize selbst. Betrachtet man lediglich die Paper-Pencil-Tests der Aufmerksamkeit und der Konzentration, so lassen sie sich hinsichtlich ihrer theoretischen Konzeption zwar unterscheiden, auf Operationalisierungsebene allerdings sind sie nicht voneinander zu trennen.

2.5.2 Konzentration und Mental Speed

Die dem TEMEKKO zugrunde gelegte Definition von Konzentration (vgl. 2.1 Konzentration) beinhaltet, dass u.a. *schnell* gearbeitet wird. Diese Schnelligkeit spiegelt sich bereits im Wahrnehmungsprozess (perceptual speed) des Reizes wieder, der konzeptionell dem Aufmerksamkeitsaspekt zuzuordnen ist. Die Schnelligkeit bei der Markierung von Testaufgaben ist unter motorischer Schnelligkeit zu subsumieren. Diese sollte bei kognitiven Leistungstests durch einen adäquaten Antwortmodus zu vernachlässigen sein. Jene Speed-Leistung, die bei konzentrativen Aufgaben von maßgeblicher Bedeutung ist, ist die Schnelligkeit, mit der Reize verarbeitet werden (processing / mental speed).

Unter Mental Speed versteht man die Geschwindigkeit, mit der Informationen verarbeitet werden (Nettelstroth, 2003). Schweizer (1995b) gliedert die Forschung zur mentalen Informationsverarbeitung nach zwei Gesichtspunkten. (1) Der Ansatz zu den kognitiven Korrelationen befasst sich mit elementaren Leistungen, die zur mentalen Informationsverarbeitung benötigt werden. (2) Der Ansatz der kognitiven Komponenten beschäftigt sich mit den Anforderungen, die Leistungstestaufgaben an die Probanden stellen. Ausgangspunkt der Forschung zu kognitiven Korrelaten der Intelligenz war die Geschwindigkeitshypothese. Sie konstatiert, dass Unterschiede individueller Intelligenzleistungen auf die Unterschiedlichkeit der Geschwindigkeit der Informationsverarbeitung zurückzuführen sind (Neubauer, 1993). Die Untersuchung der Hypothese brachte in diesem Zusammenhang Erkenntnisse über die Organisation des Gedächtnisspeichers, der Aufmerksamkeit sowie über Mental Speed, und sie wurden als Korrelate der Intelligenz mit Intelligenztestleistungen in Verbindung gebracht (Bühner, 2001).

Die Verfahren zur Erfassung von Mental Speed verwenden unterschiedliche Reaktions- und Wahlreaktionsaufgaben, sowie verschiedene experimentalpsychologische Paradigmen. Nachfolgend soll lediglich auf das Inspection-Time-Paradigma (Vickers, Nettelbeck & Wilson, 1972) stellvertretend eingegangen werden. Die Inspection Time (Vickers & McDowell, 1996) ist jene „geschätzte Zeit, die ein Individuum benötigt, um eine einzelne sensorische Beobachtung zu machen“ (vgl. Bühner, 2001, S. 46). Einzelne Reize werden den Probanden kurz (10 bis mehrere 100 Millisekunden) dargeboten, auf die instruktionsgemäß nach einer einfachen Regel ohne zeitliche Begrenzung zu reagieren ist. Die Zeit, die benötigt wird, um auf den Reiz richtig zu reagieren, wird als Stimulus-Aussetzungsdauer erfasst. Als Kennwerte für Mental Speed gelten Tempoleistungen und Reaktionszeiten (Bühner, 2001). Der Zahlen-Verbindungs-Test (ZVT), (Oswald & Roth, 1987) dient der Erfassung der Informationsverarbeitungsgeschwindigkeit im o.g. Sinn und kann als Mental Speed Test angesehen werden.

Zusammenhänge zwischen Mental Speed Tests und anderen Tests, auf denen ebenso schnell auf einfache Reize zu reagieren ist (z.B. Konzentrationstests oder Aufmerksamkeitstests) liegen unterschiedlich hoch (Bühner, 2001). Oswald und Roth (1987) berichten eine Korrelation zwischen Tempoleistungen im Test d2 (Brickenkamp, 2002) und dem ZVT von $r = |.48|$. Schweizer & Moosbrugger (1999) zeigen eine Korrelation zwischen der Gesamtzahl bearbeiteter Zeichen im Frankfurter Aufmerksamkeits-Inventar FAIR (Moosbrugger & Oehlschlägel, 1996) und der Leistung in einem von Schweizer (1996) entwickelten Mental-Speed-Test von $r = .40$. Die Subtests des Berliner-Intelligenz-Strukturtests zur Bearbeitungsgeschwindigkeit (BIS-B) korrelieren mit Aufgaben zur Erfassung von Mental Speed (sog. elementary cognitive tasks) bei Jäger, Süß und Beauducel (1997) um $r = .75$. Die Korrelationen von Mental-Speed-Tests und

Intelligenztests werden nach Amelang (1996) als gering bzw. nach Schweizer (1995a) als mittel bewertet.

Die Geschwindigkeit der Informationsverarbeitung (mental oder processing speed) kann als bedeutende Basis bei der Erbringung kognitiver Leistungen angesehen werden (Rindermann & Neubauer, 2004). Insbesondere bei sehr einfachen kognitiven Aufgaben gewinnt die Bearbeitungsgeschwindigkeit an Bedeutung. In den zur Validierung des TEMEKKO verwendeten Stichproben wurde aber kein Verfahren verwendet, das typischerweise explizit Informationsverarbeitungsgeschwindigkeit (mental speed) erfasst (z.B. ZVT Zahlen-Verbindungs-Test). Es sei daher lediglich darauf hingewiesen, dass Mental Speed (wahrscheinlich) am deutlichsten in den verwendeten Aufmerksamkeitstests zum Tragen kommt. Aufmerksamkeit definiert sich als Selektion in der Wahrnehmung und trägt daher die Wahrnehmungsgeschwindigkeit (perceptual speed) als Teil der Bearbeitungsgeschwindigkeit in sich.

Die von Marschner (1980) im Manual des Revisionstests berichteten Korrelationen mit den Subtests 9 und 10 des Prüfsystems zur Schul- und Bildungsberatung (PSB) von $r = .54$ und $r = .60$ werden als Konvergenz der Konstrukte Konzentration und Wahrnehmungsgeschwindigkeit bewertet (Bühner, 2001), wenngleich der PSB als Intelligenz-Diagnostikum geführt wird. Die Wahrnehmungsgeschwindigkeit wird etwa durch Aufgabenstellungen operationalisiert, die den Durchstreichetests zur Erfassung der Konzentration formal ähnlich sind (Bühner, 2001).

2.5.3 Konzentration und Intelligenz

Das Modell, das alle bislang entwickelten Aufgaben in Intelligenztests sowie die Strukturen der einzelnen Modelle in sich integriert, ist das Modell des Berliner Intelligenzstruktur-Tests (Jäger, Süß & Beauducel, 197). Das Modell ist hierarchisch und bimodal in seiner Struktur aufgebaut und ist Bestandteil zahlreicher Forschungsarbeiten (z.B. Carroll, 1993; Gustafsson, 1988; Neubauer & Bucik, 1996; Süß, 1995; Wittmann, 1987, 1988). Das Modell erfasst Intelligenz auf einer sehr breiten Basis nach zweierlei Aspekten, i.e. nach operativen Fähigkeiten (Verarbeitungskapazität, Einfallsreichtum, Bearbeitungsgeschwindigkeit und Merkfähigkeit) und inhaltsgebundenen Fähigkeiten (sprachgebundenes Denken, zahlengebundenes Denken sowie figural-bildhaftes Denken). Bearbeitungsgeschwindigkeit wird von Jäger et al. (1997, S. 6) als „Arbeitstempo, Auffassungsleichtigkeit und Konzentrationskraft beim Lösen einfach strukturierter Aufgaben mit geringem Schwierigkeitsniveau“ verstanden. Diese Definition lässt Zusammenhänge mit dem dem TEMEKKO zugrunde liegenden Konzentrationsbegriff erkennen. Die Fähigkeiten der einzelnen Facetten weisen zueinander moderate, aber durchwegs positive Zusammenhänge auf, was auf ihre Aggregierbarkeit hindeutet. Hierarchisch darüber steht die allgemeine Intelligenz als Integral aller Fähigkeiten.

Besonders hervorzuheben ist der Zahlen-Symbol-Test, der als figuraler Subtest im BIS-Modell der Bearbeitungsgeschwindigkeit zugeordnet ist. Dieser Test ist ebenfalls Teil des Hamburg-Wechsler-Intelligenztests für Erwachsene (HAWIE-R) (Tewes, 1991). Dort gilt er als Subtest, der Konzentration erfasst. Diese Beschreibung deckt sich mit Befunden über moderate bis hohe Korrelationen zu gebräuchlichen Konzentrationstests (Bühner, 2001; Westhoff & Kluck, 1993). So berichtet Marschner (1980) von einem Zusammenhang des Zahlen-Symbol-Tests mit dem Revisions-Test (Rev.-T.) von $r = .72$ und Brickenkamp

(2002) von $r = .68$ für die Korrelation des Tempowerts im Test d2 und der Leistung im Zahlen-Symbol-Test. Enders (2004) berichtet, dass Tests zur Erfassung der Verarbeitungskapazität und von zahlengebundenen Anforderungen (nach dem BIS-Modell) sich nicht von Konzentrationstests, die das Erkennen und Bewerten von Relationen als Anforderung aufweisen, auf Operationalisierungsebene abgrenzen lassen.

2.5.4 Konzentration und Arbeitsgedächtnis

Die dem TEMEKKO zugrunde gelegte Definition (vgl. 2.1 Konzentration) einer konzentrativen Leistung sieht das schnelle und genaue *Arbeiten* unter erschwerten Bedingungen vor. Eine Aufgabe verlangt die Manipulation der Information unter erschwerten Bedingungen. Die Leistung geht durch die Bearbeitung der Reize über eine simple Reizwiedererkennung hinaus und erfordert darüber hinaus die Unterdrückung von Störreizen.

Seit dem ersten Ursprungsmodell des Arbeitsgedächtnisses von Baddeley & Hitch (1974) hat sich die Forschung um das Konzept Arbeitsgedächtnis weiterentwickelt, differenziert und spezifiziert. Gleichwohl existiert gegenwärtig keine allgemein anerkannte Definition, aber eine Vielzahl von Modellen. Die Arbeit zum Modell des Arbeitsgedächtnisses von Baddeley (1986) ist wohl eine der bekanntesten Arbeiten über kognitive Inhalte, die sich mit Exekutiven Funktionen befasst. Die begriffliche und inhaltliche Verwobenheit zwischen Arbeitsgedächtnis und exekutiven Funktionen wurde bereits in Abschnitt 2.4 (Exekutive Funktionen) deutlich. Das Modell von Baddeley umfasst bekanntermaßen zwei Subsysteme (phonologische Schleife sowie räumlich-visueller Notizblock) und die zentrale Exekutive, die kognitive Prozesse kontrolliert und steuert und demnach für exekutive

Funktionen verantwortlich ist. Je nach Modell (und Forschung) werden einzelne exekutive Funktionen als mehr oder weniger relevant für die Steuerung von Gedächtnisinhalten angesehen. Der zentralen Exekutive bei Baddeley entspricht das Supervisory Attentional System (SAS) bei Norman & Shallice (1986) oder das Konzept der kontrollierten Aufmerksamkeit (controlled attention) bei Engle, Tuholski, Laughlin & Conway (1999). Nach Engle et al. (1999) stellt die kontrollierte Aufmerksamkeit die entscheidende Komponente des Arbeitsgedächtnisses dar; eine domänenunabhängige Aufmerksamkeitskapazität, um Repräsentationen im Arbeitsgedächtnis aktiv aufrecht zu erhalten oder zu unterdrücken. Darüber hinaus zeigen Befunde von Luciano, Wright, Smith, Geffen, Geffen und Martin (2001), dass die Arbeitsgedächtniskapazität und Intelligenz wiederum gemeinsame Varianz aufweisen. Dies unterstreicht die enge Verwobenheit der Konstrukte und wie naheliegend – wenngleich auch verwirrend – es ist, Konzepte, die unterschiedlichen Konstrukten zugeordnet werden, inhaltlich mehrdeutig zu verwenden.

Zur Validierung des TEMEKKO wurde das Arbeitsgedächtnismodell nach Oberauer, Süß, Wilhelm und Wittmann (2003) einbezogen. Das Modell wurde auf einer breiten Basis von Aufgaben und Paradigmen auf faktorenanalytischer Basis entwickelt. Ausgehend von der Untersuchung zur Binnenstruktur der Arbeitsgedächtniskapazität durch Oberauer, Süß, Wilhelm & Wittmann (1999) und Süß, Oberauer, Wittmann, Wilhelm & Schulze (1996) und des Zusammenhangs von Arbeitsgedächtnis und Intelligenz entwickelten Oberauer et al. (2000, 2003) ein Modell, das eine Differenzierung mit zwei Facetten – analog zum BIS-Modell der Intelligenz - darstellt. Eine Facette beschreibt die Funktionen der Arbeitsgedächtnisressourcen (simultanes Speichern im Kontext gleichzeitiger Verarbeitung, Supervision und Coordination bzw. Relational Integration), die andere Facette die Inhalte des Aufgabenmaterials (verbal-numerisch und figural-räumlich). Speichern im Kontext gleichzeitiger Verarbeitung definieren Oberauer et al. (2003) als die

Fähigkeit, kurzzeitig dargebotene Informationen über einen gewissen Zeitraum aufrecht zu erhalten, während dessen sie nicht mehr verfügbar ist. Die Verarbeitung der Information wird als Transformation oder Ableitung neuer Information – nicht als die simple Wiedergabe oder Gruppierung von Items – verstanden. Diese Fähigkeiten werden über Dual-Task Aufgaben (eine Spannen-Aufgabe für die Speicherkomponente und eine Wahlreaktions-Aufgabe der selben Inhaltsdomäne für die Verarbeitungskomponente) erfasst. Supervision (shifting) kann als Überwachung laufender kognitiver Prozesse und Handlungen sowie die selektive Aktivierung von relevanten Repräsentationen und die Unterdrückung irrelevanter, ablenkender Informationen gesehen werden. Sie werden typischerweise über Set-Switching-Aufgaben nach Rogers & Monsell (1995) erfasst, wobei spezifische und generelle switching Kosten ausgewertet werden. Die Koordinationsfunktion (Coordination) bezieht sich auf die Fähigkeit, Informationen in neue Beziehungen zu einander zu bringen und diese in übergeordnete Strukturen zu integrieren (Waltz, Knowlton, Holyoak, Boone, Mishkin & de Menezes Santos, 1999). Diese Anforderungen von Coordination werden präziser mit dem Begriff Relational Integration beschrieben, wie von Oberauer, Süß, Wilhelm und Wittmann (2008) vorgeschlagen. Aufgaben zur Erfassung von Coordination bzw. Relational Integration sind Überwachungsaufgaben (monitoring) in den verschiedenen Inhaltsdomänen, wie z.B. Flugüberwachung (figural).

Bühner, Krumm, Ziegler und Plücken (2006) fanden den höchsten Zusammenhang auf latenter Ebene zwischen den Komponenten der Arbeitsgedächtniskapazität (Speichern im Kontext gleichzeitiger Verarbeitung, Supervision und Coordination) sowie Konzentration (sustained attention) zwischen Coordination und Konzentration ($r = .77$). Darüber hinaus konnten sie zeigen, dass zur Vorhersage von Reasoning die Komponenten der Arbeitsgedächtniskapazität und Konzentration unterschiedlich beitrugen. So konnten

Coordination und Speichern im Kontext gleichzeitiger Verarbeitung maßgeblich zum erklärbaren Varianzanteil beitragen, wohingegen Supervision und Konzentration dies nicht gelang. Konzentration wurde in der Untersuchung als Sustained Attention benannt und über die Tests d2 und FAKT operationalisiert. Vertiefend wurde berichtet, dass Konzentration keine Prädiktion an Intelligenz unabhängig von Coordination leistete.

Der nun folgende Abschnitt befasst sich mit der Entwicklung des TEMEKKO und insbesondere mit seiner Aufgabenentwicklung im Detail. Vertieft wird diese Darlegung mit Erläuterungen zur Aufgabenstellung, Fehlerarten sowie der Kontrolle von instruktionswidriger Bearbeitung der Testaufgaben. Hieran schließen sich die Gestaltung der Testformen sowie Beschreibungen zur Testdurchführung und -auswertung an.

3 Entwicklung des TEMEKKO

Für die Konstruktion des TEMEKKO wurde ein top-down Ansatz gewählt. Ausgehend vom Merkmalsbereich Konzentration wurde zunächst versucht, diesen anhand von in der Literatur wiederkehrenden Charakteristika einzugrenzen. Die für den TEMEKKO zu Grunde gelegte Definition besagt: "Konzentration ist die Fähigkeit, unter Bedingungen, die das Erbringen einer kognitiven Leistung normalerweise erschweren, schnell und genau zu arbeiten.“ (Schmidt-Atzert, Büttner, Bühner, 2004; Schmidt-Atzert, Krumm, Bühner, 2008) (vgl. 2.1 Konzentration). Basierend auf dieser Definition und dem Modell zu Konzentration (und Aufmerksamkeit) nach Schmidt-Atzert, Krumm und Bühner (2008) (vgl. 2.5.1 Konzentration und Aufmerksamkeit) wurde anschließend nach rationaler (deduktiver) Methode versucht, Aufgaben zu generieren, die als objektive Indikatoren die Fähigkeitsausprägung erfassen.

3.1 Itemkonstruktion

Die Konstruktion der Aufgaben des TEMEKKO basiert zum einen auf den Anforderungen, die ein Konzentrationstest erfüllen muss (vgl. 2.2 Konzentrationstests), und zum anderen auf theoretischen Überlegungen zu Aufgabenanforderungen, die maximal unterschiedlich zu divergenten Konstrukten sein sollten. Die von Bühner (2001) dargestellte Zusammenfassung aller wesentlichen Merkmalskriterien von Konzentrationstests sind im TEMEKKO direkt operationalisiert worden:

1. aktive intentionale Zuwendung auf Reize (Westhoff 1995; Berg & Imhof 1998, vgl. auch Konzentration als Willensakt von Mierke 1957)

Der TEMEKKO verlangt vom Probanden durch die Aufgabenstellung, eindeutig definierte Zielobjekte zu finden und zu markieren, beansprucht somit eine intentionale Zuwendung auf Reize.

2. fortlaufende Reizselektion (Westhoff 1991, 1995)

Der TEMEKKO verlangt durch die Bearbeitung von insgesamt 1.140 Items nacheinander (190 Items pro Materialart (figural, numerisch, verbal) und Testform (zwei Testformen A und B), fortlaufend Reize zu selektieren.

3. aufmerksame Wahrnehmung von Reizen (Westhoff 1991, 1995; Hoffmann 1993)

Der TEMEKKO erfordert eine achtsame Wahrnehmung der Reize, um Zielreize unter Distraktoren auszuwählen.

4. kontrollierte Verarbeitung von Reizen (Berg & Imhof, 1998)

Der TEMEKKO erfordert die Verarbeitung von fortwährend wechselnden Reizen gezielt unter der Maßgabe, dass die Relationen zwischen den Item-Elementen im Sinne der Aufgabenlösung analysiert werden müssen.

5. Abschirmung von Störungen bei der Bearbeitung einer Aufgabe (Berg & Imhof 1998, Neumann & Simon 1994)

Der TEMEKKO erfordert das Ausblenden irrelevanter Zusatzmerkmale der zu bearbeitenden Reize. Durch unterschiedliche Größe/Breite, Position, Druckbild oder Abstände der Item-Elemente wird die Verarbeitung der Reize deutlich erschwert.

6. Beantwortung eines Reizes unter handlungsführenden Gesichtspunkten (Posner & Rafal 1987)

Der TEMEKKO verlangt bei Entdeckung eines Zielreizes eine instruktionsgemäße Handlung (Markieren des Targets).

7. Beanspruchung energetischer Ressourcen (Berg & Imhof 1998)

Der TEMEKKO verlangt die Aufrechterhaltung der Fokussierung über den gesamten Testungszeitraum von 14 Minuten Netto-Bearbeitungszeit hinweg.

3.1.1 Items und Aufgabe

Ein Item setzt sich jeweils aus drei Elementen in einem einheitlich großen Rahmen zusammen. Elemente sind Balken im figuralen Bereich, Ziffern im numerischen Teil und Wörter für das verbale Material. Die Aufgabe besteht darin festzustellen, ob die drei Item-Elemente zueinander eine auf- oder absteigende Reihe bilden oder nicht. Im figuralen Teil ist die Höhe der Balken entscheidend, die numerischen Item-Elemente werden nach ihrem Zahlenwert geprüft. Die Entscheidung der verbalen Items ist nach der semantischen Bedeutung zu treffen. Ein Beispiel für eine aufsteigende (Zahlen-)Reihe ist „1 3 5“, ein Beispiel für eine abfallende „7 5 1“. Es gibt demnach zwei Arten von Zielobjekten, die beide gleich zu beantworten, nämlich durchzustreichen, sind. Somit wurde ein gebundenes Antwortformat gewählt und es gibt jeweils eine eindeutig richtige oder falsche Lösung. Distraktoren sind Elemente, die keine auf- oder absteigenden Reihen darstellen. Ein Beispiel ist „1 5 3“.

Die Zahl der möglichen Targets und Distraktoren ist groß genug, um ein Wiedererkennen von bereits bearbeiteten Items durch Erinnerung zu verhindern. In der figuralen Version ist es besonders wichtig, dass die drei Elemente eines Items visuell gut unterscheidbar sind. Aus diesem Grund konnten nicht alle Kombinationen realisiert werden. Die figuralen Item-Elemente unterscheiden sich immer um mindestens 2 mm in der Höhe. Deshalb kommt etwa die Zahlenkombination „1 2 3“ (analog: 1 mm, 2 mm, 3 mm) nicht vor. Nach einer „1“ stellt die „3“ die nächste erlaubte Größe dar. Auf die „3“ darf keine „4“ oder „2“ folgen, sondern entweder wieder eine „3“, oder eine „5“ bzw. noch größere Zahl oder eine „1“. Insgesamt sind damit 35 auf- und 35 absteigende Reihen möglich, die dieser Anforderung genügen (vgl. Anhang V). Als Distraktoren sind entsprechend 347 Kombinationen möglich (vgl. Anhang V). Für eine Testseite werden 48 Targets (ca. 25 %) und 142 Distraktoren aus dem Pool aller möglichen Kombinationen über 19 Zeilen verwendet. Für jede Materialart ist pro Testform eine Testseite vorgesehen. Für die Instruktion und Beispiele sowie Übungsaufgaben dienen Items, die in den Testseiten nicht vorkommen. Damit ist sichergestellt, dass bei jedem Item erneut jene kognitiven und exekutiven Prozesse ablaufen, die bei seiner Lösung erforderlich sind.

Um Ablenkung zu erzeugen, wurden einzelne Elemente mit zusätzlichen Merkmalen versehen, die für die Aufgabenbearbeitung irrelevante Reize darstellen. Solche Zusatzmerkmale sind das Druckbild, die Größe/Breite, die Position, sowie der Abstand der Item-Elemente zueinander. Bei unterschiedlicher Druckqualität sind die Zahlen oder Wörter fett oder normal gedruckt. Im figuralen Teil sind die Balken analog zum Fettdruck der beiden anderen Materialarten ausgefüllt. Unterschiedliche Schriftgröße bedeutet, dass Zahlen oder Wörter normal groß (Arial 20) oder kleiner (Arial 14) gedruckt sind. Diese Besonderheit stellt sich im figuralen Teil als reduzierte Balkenbreite (4 mm versus 2 mm) dar. Als drittes Zusatzmerkmal sind die Zahlen oder Balken entweder in der horizontalen

Position hochgestellt oder nicht. Als analoges Zusatzmerkmal im verbalen Teil sind die Wörter an den linken Rand des Kästchens gezogen. Darüber hinaus wird der Abstand zwischen zwei Item-Elementen variiert (normal versus vergrößert). Die Verteilung der Zusatzmerkmale über die Item-Elemente erfolgte nach dem Zufallsprinzip. Nichtsdestotrotz wurden die Häufigkeiten der verwendeten Zusatzmerkmale über die beiden Testformen (A und B) hinweg gleich gehalten. Da zugleich mehrere Zusatzmerkmale eines Item-Elements verändert sein können, ergeben sich acht verschiedene Darstellungsmöglichkeiten pro Item-Element, i.e. 1.536 Varianten eines einzigen Items.

Für die gezielte Veränderung der Item-Elemente wurden die einzelnen Zusatzmerkmale mit einer genau bestimmbar Häufigkeit verändert. 7 von 48 Zielreizen (14.6 %) haben keine Zusatzmerkmale. Bei den 41 übrigen Targets wurden die Zusatzmerkmale systematisch ergänzt. Diese Systematik der Targets wurde auf die Gesamtzahl der Distraktoren ($n = 142$) hochgerechnet, sodass die relativen Häufigkeiten gleich sind (20 Distraktoren ohne Zusatzmerkmale, i.e. 14.1 %). Welche Items unverändert blieben, und welche Elemente wie verändert wurden (Qualität und Quantität der Zusatzmerkmale) erfolgte nach dem Zufallsprinzip. Ein Item-Element konnte in mehr als nur in einem Merkmal verändert werden, oder auch gar nicht.

Das Testformular besteht aus 19 Zeilen mit je 10 Items. Das Verhältnis von Targets zu Distraktoren liegt bei 1:3. Insgesamt werden 190 Items pro Inhaltsdomäne und Testform dargeboten, von denen 48 Targets und demnach 142 Distraktoren sind. Die Abstände zwischen den Targets variieren zufällig, damit keine Erwartungen hinsichtlich des nächsten Zielreizes gebildet werden können. Darüber hinaus wurden 7 mögliche Abstände (0 bis 6) zwischen zwei aufeinanderfolgenden Targets festgelegt. Jeder Abstand kommt

mit einer relativen Häufigkeit von $1/7$ vor. Die genaue Abfolge der Abstandsvarianten (0 bis 6) erfolgte nach dem Zufallsprinzip. Eine Einschränkung besteht lediglich insoweit, als dass bis zur Wiederholung eines Abstandes mindestens 3 andere Varianten vorkommen mussten. Dieses Prinzip führte dazu, dass pro Zeile 1 bis 3 Targets vorhanden sind. Anhang V zeigt die Verteilung der Items. Nach dieser grundsätzlichen Festlegung der Anordnung wurde festgelegt, welche Targets bzw. welche Distraktoren besonders leicht bzw. besonders schwer sein sollten. Hierfür wurden die o.g. Überlegungen der Wirksamkeit der Itemdarstellung sowie das Bestreben einer annähernden Gleichverteilung über den gesamten Test berücksichtigt.

Ziel bei der Entwicklung des Verfahrens war es auch, die aus der Intelligenzforschung bewährte ausgewogene Berücksichtigung der Domänen - wie etwa im BIS Modell dargestellt - abzudecken. Entsprechend dieser Forderung wurden separate Versionen mit figuralem, numerischem und verbalem Material angefertigt. Die Item-Elemente bestehen entweder aus figuralem (Balken), numerischem (Ziffern) oder verbalem (Wörter) Material. Die drei Materialarten wurden nach einer strikten Übersetzungsregel ineinander überführt. Ausgehend vom numerischen Material konnte die Übersetzung leicht in das figurale Material übertragen werden, indem der Zahlenwert jeweils die Höhe der Balken in Millimeter darstellte. Das numerische Item „1 8 3“ entspricht demnach einem figuralem Item mit den Balken in den Höhen 1 mm, 8 mm und 3 mm. Bei der Übersetzung in die verbale Form wurde beachtet, dass es sich um allgemein bekannte Begriffe handelt. Die Bedeutung der Wörter musste unmissverständlich und die Größe des beschriebenen Objekts eindeutig sein. Die Ziffern wurden wie folgt übersetzt: 1 = Laus, 2 = Nuss, 3 = Hut, 4 = Stuhl, 5 = Mann, 6 = Pferd, 7 = Haus, 8 = Berg, 9 = Welt. Hieraus ist ersichtlich, dass der Anstieg des Zahlenwerts einher geht mit dem Anstieg der Größe des verbal benannten Objekts/Lebewesens.

Das Verfahren wurde in zwei Testformen entwickelt (A und B). Die Testform B wurde aus den Aufgabenzeilen der Testform A gebildet. In Tabelle 1 wird die analoge Übertragung der Testzeilen dargestellt.

Tabelle 1: Analogie der Testzeilen der Formen A und B

Zeile der Testform A	Zeile der Testform B
1	10
2	11
3	12
4	13
5	14
6	15
7	16
8	17
9	18
10	19
11	1
12	2
13	3
14	4
15	5
16	6
17	7
18	8
19	9

Es gilt zu beachten, dass die Item-Elemente der Testform A in umgekehrter Reihenfolge in die Testform B übernommen wurden. So wurde aus einem Item " 8 6 4" in der Form A das Item "4 6 8" in der Form B. Die Zusatzmerkmale blieben dahingehend auf der jeweiligen Position des Item-Elements bestehen, sodass sich Erscheinungsbilder, wie in Abbildung 2 dargestellt, für analoge Items ergeben können. Gleiche Item-Elemente, z.B. " 7 7 7" wurden insofern übertragen, als dass die Relation zum jeweils gegenläufigen Randelement gleich blieb, i.e. " 3 3 3".



Abbildung 2: Item A 156 und analoges Item B 56

3.1.2 Fehlerarten

Die Items sind so beschaffen, dass sowohl Auslassungs- als auch Verwechslungsfehler in nennenswerter Häufigkeit vorkommen. Bei einem Teil der Distraktoren - sog. "besonders schwere Distraktoren" - ($n = 17$) werden *Verwechslungsfehler* provoziert, indem die Elemente durch ihre optische Erscheinung als Teil einer auf- oder absteigenden Reihe wirken. Hierzu sind insbesondere Kombinationen geeignet, die nicht stark von einer Reihe abweichen (z.B. „3 7 5“ oder „9 2 4“). Die Hochstellung eines Elements suggeriert, dass die Reihe an dieser Stelle aufsteigt, wenngleich auch der Zahlenwert gering ist. Durch optische Verkleinerung eines Elementes wird der Eindruck erweckt, dass die Reihe an dieser Stelle abfällt, wenngleich auch der Zahlenwert hoch ist. Zusätzlich wird durch ungewohnte Reize (Fettdruck oder Abstand vergrößert) die Verarbeitung des Items erschwert. Es ist davon auszugehen, dass die veränderte Darstellung der Item-Elemente – als interferierende Information – die Probanden von der zu verarbeitenden Information ablenkt.

Bei einem Teil der Zielobjekte - sog. "besonders schwere Targets" ($n = 10$) wird ein Erkennen als Reihe erschwert, indem *Auslassungsfehler* provoziert werden. Als solche eignen sich Zielreize insbesondere, wenn:

- zwei Elemente nahe beieinander liegen (nur zwei Schritte Abstand) oder das Item kein Randelement (1 oder 9) enthält, bei dem vermutlich leicht zu erkennen ist, dass es kleiner (1) oder größer (9) als die übrigen Elemente ist.
- die Items an unerwarteter Position stehen (kurz oder unmittelbar nach einem anderen Zielreiz oder als allererstes bzw. allerletztes Target einer Testform).

Über den gesamten Test hinweg sollte die Zahl der Auslassungs- und der Verwechslungsfehler konstruktionsgemäß deutlich von Null verschieden sein und interindividuell variieren.

3.1.3 Kontrolle von instruktionswidrigem Bearbeiten

Generell besteht bei Testungen die Gefahr, dass Probanden versuchen, ihre Testwerte zu verfälschen. Hierbei unterscheiden Bortz und Döring (2002, S. 230) danach, ob versucht wird, möglichst hohe Testwerte zu erzielen (Simulation) oder möglichst niedrige (Dissimulation). Beide Möglichkeiten sind unterschiedlich motiviert, nichtsdestotrotz werden sie angewandt, wenn die getestete Person sich aus verfälschten Werten einen Vorteil verspricht (Griffith, Chmielowski & Yoshita, 2007). Es gilt für objektive Leistungstests gleichermaßen, dass Personen möglicherweise die Items nicht ernsthaft bearbeiten, sondern entweder zufällig durchstreichen, oder Items überspringen, um ihre Mengenleistung zu steigern (Berg & Erlwein, 1991; Oehlschlägel & Moosbrugger, 1991). Dieser Gefahr kann die Verwendung des KL-Wertes (vgl. 2.3.1 Kennwerte der Konzentrationsleistung) als Leistungsmaß entgegen wirken, sodass ein Überspringen von Zeichen oder ganzen Testzeilen das individuelle Leistungsmaß nicht fälschlicherweise zu hoch einstuft. Wenngleich auch Schmidt-Atzert, Bühner, Rischen und Warkentin (2004) in einer Untersuchung am Test d2 zeigen, dass Probanden eine Simulation mittels Ü-Syndrom praktisch so gut wie nicht anwenden. Wahrscheinlich sind objektive Leistungstests nur nach unten verfälschbar (Ziegler, Schmidt-Atzert, Bühner & Krumm, 2007), i.S. einer Dissimulation. Allerdings übertreiben Probanden meist, wenn sie absichtlich schlechtere Testleistungen erzielen wollen, dass selbst hirngeschädigten Patienten normalerweise weniger Fehler unterlaufen (Schmidt-Atzert, Bühner, Rischen &

Warkentin, 2004). Eine Kontrollmöglichkeit hierzu stellt die Prüfung von besonders leichten Items dar (Schmidt-Atzert & Bühner, 1998). Einige Items sind eindeutig als Targets bzw. Distraktoren erkennbar. Normalerweise sollten bei diesen Items keine Fehler gemacht werden. Sie können dazu dienen, Bearbeitungsstrategien wie Raten, willkürliches Anstreichen (Dissimulation), oder fehlendes Instruktionsverständnis aufzudecken. Die Platzierung der sog. "besonders leichten Items" ist in den Tabellen V.1 und V.2 des Anhangs V für die jeweiligen Testformen aufgezeigt. In diesem Zusammenhang sei auf Kapitel 3.4 (Auswertung) hingewiesen, in dem Grenzen benannt werden, nach denen ein Testbogen als "nicht auswertbar" eingestuft werden muss.

Als "besonders leichte Targets" ($n = 9$) werden Items verstanden, wenn:

- die Elemente weit voneinander entfernt sind und mindestens ein Element eine Randposition hat (1 oder 9), z.B. „1 4 8“ und
- die Elemente sich in ihren Zusatzmerkmalen nicht unterscheiden, bzw. keine Zusatzmerkmale haben.

Als "besonders leichte Distraktoren" ($n = 30$) werden Items verstanden

- mit 3 identischen Elementen (z.B. „1 1 1“),
- wenn sie sich nicht in ihren Zusatzmerkmalen unterscheiden oder
- die eine symmetrische Anordnung aufweisen, z.B. „1 9 1“, bei denen das mittlere Element sich von den beiden anderen mittels Zusatzmerkmalen (Fettdruck oder vergrößerter Abstand) unterscheidet.

3.2 Gestaltung des Testformulars

Das Deckblatt des Testformulars enthält allgemeine Fragen zur Person, wobei der Name des Probanden zur Wahrung der Anonymität auch als Kodierung notiert werden kann. Wichtig ist lediglich, dass bei Durchführung verschiedener Tests der Proband immer den selben Namen / die selbe Kodierung benutzt. Die übrigen Fragen umfassen u.a. Alter und Geschlecht, Bildungsstand, Händigkeit und Sehhilfe bzw. ob eine Form der Beeinträchtigung vorliegt (physisch oder aufgrund von Medikamenten).

Nach Aufblättern des Testformulars ist auf der rechten Seite die allgemeine Instruktion zum Verfahren gedruckt, sowie einige erste Beispielitems. Die nächste umgeblätterte Seite beinhaltet die Instruktion, spezifisch zum ersten Testteil, zum figuralen Material mit einigen Beispielitems und zwei Zeilen Übungsaufgaben. Auf den nächsten beiden Seiten befinden sich auf der linken Seite die Aufgaben der Testform A und auf der rechten Seite die Aufgaben der Testform B. Durch weiteres Umblättern gelangt man zur spezifischen Instruktion für den nächsten Testteil, das numerische Material, analog aufgebaut zum figuralen Teil. Auch die beiden nächstfolgenden Testseiten enthalten analog zum figuralen Teil nun aber die beiden numerischen Testformen. Auf der linken Seite Form A und auf der rechten Seite Form B. Die übrigen Testblätter zum verbalen Material zeigen wie die beiden anderen Materialarten auch zunächst die spezifische Instruktion, gefolgt von der Testform A auf der linken und der Testform B auf der rechten Seite.

In der Kopfzeile jedes Aufgaben-Testblattes wird die Aufgabenstellung gut sichtbar wiederholt „Alle Balkenreihen durchstreichen!“ bzw. „Alle Zahlenreihen durchstreichen!“ bzw. „Alle Wortreihen durchstreichen!“. Am Beginn jeder Testzeile befindet sich links vor dem ersten zu bearbeitenden Item ein Kästchen, das die Nummerierung der in der

jeweiligen Zeile stehenden Items benennt. In der ersten Zeile steht in diesem Kästchen zusätzlich das Wort „Start“ mit einem Pfeil, der die Bearbeitungsrichtung anzeigt. Diese Kästchen sind grau unterlegt, die Nummern sind in kleiner Schrift (Arial 12) gedruckt, sodass sie sich deutlich von den Zeichen in den Item-Kästchen unterscheiden. Diese Darstellung erleichtert die Auswertung der Mengenleistung. Ist beispielsweise das letzte bearbeitete, eingekreiste Item in Zeile 12 das dritte Zeichen, so wurden 123 Items insgesamt bearbeitet.

3.3 Durchführung

Der TEMEKKO kann für Einzel- sowie Gruppentestungen angewandt werden. Die Reihenfolge der Durchführung bleibt hierbei immer gleich. Der Testleiter verteilt an die Probanden jeweils einen Testbogen, bevor er die Aufgabenstellung näher erläutert. Zunächst wird die Instruktion vom Testleiter laut vorgelesen. Die Probanden werden angehalten, für sich mit zu lesen. Einleitend wird der Proband in der Instruktion darauf hingewiesen, dass es darauf ankommt, sich gut zu konzentrieren. Danach wird der Aufbau des Testhefts dargestellt. Hierauf folgt die Erläuterung der Aufgabenstellung: Eine Aufgabe besteht aus einem Kästchen, in dem drei Elemente (Balken, Zahlen *oder* Wörter) stehen. Es sollen jene Kästchen herausgefunden und markiert werden, deren Elemente nach inhaltlichen Gesichtspunkten eine Reihe bilden. In der figuralen Testform ist die Höhe der einzelnen Balken ausschlaggebend, der numerische Teil orientiert sich am Zahlenwert und beim verbalen Material ist der semantische Inhalt der Wörter maßgeblich. Dabei kann eine Reihe aufsteigend oder abfallend sein. Beide Arten von Zielreizen (aufsteigende sowie abfallende Reihenfolgen) werden anhand von Beispielen erläutert. Danach wird darauf hingewiesen, eindeutige Markierungen vorzunehmen (am besten einen

diagonalen Strich durch das Kästchen zu ziehen) und dass ein versehentlich markiertes Item durch einen zweiten diagonalen Strich in die entgegengesetzte Richtung (insgesamt ein X) wieder rückgängig gemacht werden kann. Danach erfolgt die Erläuterung der Zusatzmerkmale sowie der deutliche Hinweis, dass diese als irrelevante Reize bei der Aufgabenbearbeitung zu ignorieren sind. Es werden dann die Reihenfolge des Testablaufs (figural (Balken), numerisch (Zahlen), verbal (Wörter)) sowie der Hinweis auf Übungsaufgaben erläutert, die die Zusatzmerkmale enthalten. Zu guter Letzt werden die Probanden informiert, dass sie erst nach dem „Start“ Signal beginnen dürfen, dass die vorgesehene Bearbeitungszeit knapp bemessen ist, und dass bei der Aufforderung „Stopp, Kreis, nächste Seite“ der Proband die Bearbeitung der Aufgaben sofort beenden und das letzte bearbeitete Item einkreisen soll. Die Probanden werden darauf hingewiesen, dass sie erwartungsgemäß nicht alle Aufgaben in der vorgegebenen Zeit bearbeiten werden können. Gleichwohl soll so schnell und genau wie möglich gearbeitet werden. Noch einmal erfolgt der Hinweis, dass schnelles und richtiges Antworten gleich bewertet werden, jedes einzelne Kästchen bearbeitet und keines übersprungen werden soll; dies würde bei der Auswertung erkannt werden. Die Bearbeitungsrichtung entspricht der Schreib- und Leserichtung.

Nach der einführenden generellen Instruktion geht der Testleiter über in die spezifische Einführung zur Bearbeitung des ersten Testteils, des figuralen Materials. Noch einmal wird die Aufgabenstellung dargelegt und auf die Besonderheiten bei der Bearbeitung eingegangen. Mit Abbildungen der Kästen auf dem Testbogen werden 4 Beispielitems dargestellt, zwei Beispiele für Zielreize, zwei für Distraktoren. In zwei darunter liegenden Testzeilen sind vom Probanden einige Übungsbeispiele zu bearbeiten. Hierfür ist keine Zeitbegrenzung vorgesehen. Es ist vom Testleiter darauf zu achten, dass die Übungsblöcke korrekt ausgefüllt werden. Zur Kontrolle für die Probanden, ob sie den Bearbeitungsmodus

der Aufgaben verstanden haben, sind unterhalb der Übungszeilen die Nummern der Zielreize gelistet. Wiederum wird ausdrücklich darauf hingewiesen, dass schnelles *und* richtiges Antworten gleichermaßen bewertet werden, bevor noch einmal das Prozedere nach dem Stopp-Signal des Testleiters ausgeführt wird. Zu guter Letzt erfolgt die Anweisung zum Umblättern und zum Start nach Aufforderung.

Für die Testform A des figuralen Materials stehen den Probanden 1 Minute und 30 Sekunden zur Bearbeitung zur Verfügung, bevor der Testleiter die Bearbeitung mit dem Stopp-Signal abbricht. Erst nach Aufforderung durch den Testleiter wird die Form B des figuralen Materials bearbeitet. Für die beiden Testformen mit numerischen Inhalten stehen den Probanden jeweils 2 Minuten zur Bearbeitung zur Verfügung. Das Start- und Stopp-Prozedere wiederholt sich analog zur ersten Materialart. Für die Formen A und B mit verbalem Material haben die Probanden 3 Minuten und 30 Sekunden Zeit, wieder eingebettet in das nun bereits bekannte Start- und Stopp-Prozedere. Die Bearbeitungszeit pro Testform und Materialart wird vom Versuchsleiter anhand einer Stoppuhr gemessen. Nachfolgende Tabelle 2 zeigt die zur Verfügung stehenden Bearbeitungszeiten im Überblick.

Tabelle 2: Bearbeitungszeiten der Materialarten und Testformen

Testform	Bearbeitungszeit
figural, Form A	1 Minute 30 Sekunden
figural, Form B	1 Minute 30 Sekunden
numerisch, Form A	2 Minuten
numerisch, Form B	2 Minuten
verbal, Form A	3 Minuten 30 Sekunden
verbal, Form B	3 Minuten 30 Sekunden

3.4 Auswertung

Wie bereits in Abschnitt 2.3.1 (Kennwerte der Konzentrationsleistung) erläutert, können als manifeste Variablen zur Beurteilung konzentrierten Arbeitens generell zum einen die verarbeitete Reizmenge und zum anderen die Zahl der dabei gemachten Fehler angesehen werden (Brickenkamp, 2002; Brickenkamp & Karl, 1986; Debus, 1995; Westhoff, 1995). Die Anzahl bearbeiteter Items pro Zeiteinheit kann als Schnelligkeitsmaß (Leistungsmaß) aufgefasst werden (Bühner, 2001; Westhoff, 1995). Als Zeiteinheit gelten die Zeitbegrenzungen pro Materialart und Testform.

Für die Formen A und B sind jeweils Auswertungsschablonen vorhanden, mit Hilfe derer der Konzentrationsleistungswert (KL-Wert), Auslassungs- und Verwechslungsfehler eruiert werden. Wie in 2.3.1 (Kennwerte der Konzentrationsleistung) erläutert, werden für den TEMEKKO die Anzahl der richtigerweise durchgestrichenen Zielreize (KL-Wert) sowie der Fehler-Prozent-Wert als maßgebliche Leistungsmaße ermittelt.

Der Konzentrationsleistungswert (KL-Wert) entspricht der Anzahl der richtig durchgestrichenen Targets. Es ist die Anzahl von Markierungen auf der Auswertungsschablone „Auslassungen“. Maximal ist ein KL-Wert von 48 pro Materialart und Testform zu erzielen.

Der GZ-Wert (Brickenkamp, 2002; Bühner, 2001; Westhoff, 1995) ist die Anzahl der bearbeiteten Zeichen einschließlich des letzten bearbeiteten (eingekreisten) Items. Wird vom Probanden verabsäumt, das letzte bearbeitete Zeichen einzukreisen, so ist die letzte Markierung auf dem Testbogen als letztes bearbeitetes Zeichen anzuerkennen. Am linken Rand der Bearbeitungszeilen lässt sich in den Kästchen die Itemnummer ablesen (z.B.

Zeile 121-130). Es müssen so nur noch die Zeichen einschließlich des eingekreisten (zuletzt markierten) Items abgezählt werden (z.B. 121, 122, 123, 124), um die Mengenleistung zu ermitteln. Es sind maximal 190 Zeichen pro Materialart und Testform zu bearbeiten. Dieser Wert dient lediglich dazu, die Auslassungen und Verwechslungen zu korrigieren. Der GZ-Wert wird nicht interpretiert.

Auslassungsfehler (AF) (Brickenkamp, 2002; Schmidt-Atzert & Bühner, 1998) sind Items, die der Proband gemäß der Instruktion markieren soll, aber fälschlicherweise nicht markiert. Hierfür ist die Anzahl der fehlenden Markierungen auf der Schablone "Auslassungen" bis einschließlich des letzten bearbeiteten/eingekreisten Items zu zählen. Es können maximal 48 Auslassungsfehler pro Materialart und Testform vorkommen, wobei dies faktisch einer Nichtbearbeitung des Tests entspricht. Werden vom Probanden mehrere Testzeilen übersprungen, so ist im Individualfall zu entscheiden, ob diese nicht bearbeiteten Zeilen als bearbeitet - mit entsprechender Anzahl von Auslassungsfehlern - anerkannt werden, oder ob das letzte markierte Zeichen - vor Überspringen der Testzeilen - als letztes bearbeitetes Zeichen zu werten ist. Ein Überspringen von Zeilen kann nur dann zur Frage stehen, wenn danach eine instruktionsgemäße Bearbeitung weiterer Zeichen augenscheinlich fortgesetzt wurde. In jedem Fall gilt eine Testform ohne Markierungen als „nicht bearbeitet“, i.e. nicht auswertbar. Die Anzahl der Auslassungen (AF) ist nach folgender Formel am GZ-Wert zu relativieren:

$$AF \% = AF / GZ * 100$$

Verwechslungsfehler (VF) (Brickenkamp, 2002; Schmidt-Atzert & Bühner, 1998) sind Zeichen, die vom Probanden gemäß der Instruktion nicht markiert werden sollen, aber fälschlicherweise markiert werden. Es ist die Anzahl der Markierungen auf der Schablone "Verwechslungen" bis einschließlich des letzten bearbeiteten/eingekreisten Items zu

zählen. Es können maximal 142 Verwechslungsfehler pro Materialart und Testform vorkommen, wobei dies faktisch einer Nichtbearbeitung i.S.v. alle Items wurden durchgestrichen entspricht. Werden vom Probanden in mehreren Testzeilen hintereinander alle Zeichen markiert, so ist im Individualfall zu entscheiden, ob diese Zeilen als bearbeitet - mit entsprechender Anzahl von Verwechslungsfehlern - anerkannt werden, oder ob das letzte nicht markierte Zeichen - vor den durchgängig markierten Zeichen/Zeilen - als letztes bearbeitetes Zeichen zu werten ist. Eine durchgängige Markierung von Zeichen/Zeilen kann nur dann zur Frage stehen, wenn danach eine instruktionsgemäße Bearbeitung weiterer Zeichen augenscheinlich fortgesetzt wurde. In jedem Fall gilt eine Testform, in der alle Items markiert sind, als "nicht bearbeitet", i.e. nicht auswertbar. Die Anzahl der Verwechslungsfehler (VF) ist nach folgender Formel am GZ-Wert zu relativieren:

$$VF \% = VF / GZ * 100$$

Zur Ermittlung des Fehlergesamtwertes in Prozent (F %) werden die Anzahl der Auslassungen (AF) und die Anzahl der Verwechslungen (VF) addiert und nach folgender Formel am GZ-Wert relativiert:

$$F \% = (AF + VF) / GZ * 100$$

In jedem Fall gilt ein Testbogen als "nicht auswertbar", wenn der Fehler-Prozent-Wert eines Probanden bei $\geq 60\%$ liegt.

Der nun folgende Abschnitt stellt die Untersuchungen zur Validierung des Tests detailliert dar. Er beschreibt insbesondere die einzelnen Schritte, die für die Analyse der Datensätze vorgenommen wurden. Für vertiefende Einblicke in die Teilergebnisse der Berechnungen sowie zu Besonderheiten bei der Analyse der einzelnen Datensets sei auf die Anhänge I bis IV verwiesen.

4 Empirische Untersuchungen

4.1 Validierungsstudien

Der TEMEKKO wurde in vier von einander unabhängigen Studien validiert. Detaillierte Beschreibungen der einzelnen Stichproben finden sich in den Anhängen I bis IV. Ein Überblick über die wichtigsten beschreibenden Charakteristika der Stichproben über die Studien hinweg ist in Tabelle 3 dargestellt.

Tabelle 3: Überblick der Charakteristika der Validierungsstichproben

Studie	<i>N</i>	Alter		Geschlecht		Anteil Studierender
		<i>M</i>	<i>SD</i>	männlich	weiblich	
1	138	69.5	7.17	35.5 %	64.5 %	0.0 %
2	144	23.7	3.37	45.8 %	54.2 %	97.9 %
3	128	21.8	2.39	24.2 %	75.8 %	98.4 %
4	121	21.1	3.25	19.8 %	80.2 %	100.0 %

4.1.1 Auswertungsmethoden

Die Auswertungen der vier Validierungsstudien erfolgten mit den Programmen SPSS für Windows (Versionen 16 und 17), AMOS (Versionen 16 und 17) sowie GPower (Version 3.0.10). Zur Unterstützung einzelner Überprüfungen der Überschreitungswahrscheinlichkeit wurde die Freeware Probability Calculator verwendet.

4.1.2 Prozedere der Datenbearbeitung

Nachfolgend wird der Ablauf der Datenbearbeitung detailliert dargestellt, wie er für jede einzelne Validierungsstudie durchgeführt wurde. Die Besonderheiten der einzelnen Datensätze, eine detaillierte Schilderung der Analysen sowie der Ergebnisse werden in den jeweils zugehörigen Anlagen I bis IV ausführlich beschrieben.

Prüfung der Daten auf Plausibilität. Zunächst wurde jede einzelne Variable der vorerhobenen Datensätze sowohl formal als auch inhaltlich auf Plausibilität geprüft. Diese Kontrolle umfasste auch jene Variablen, die nicht in nachfolgende Analysen eingingen. Die formale Prüfung beinhaltete die Prüfung auf korrektes Labelling der Daten (fehlende Werte auch als solche gelabelt; dichotome Variablen mit numerischen Werten erfasst, etc.) sowie augenscheinlich, ob Variablen, die Reaktionszeiten erfassten, realistische Werte darstellen ($RZ > 100$ ms). Die inhaltliche Prüfung umfasste eine Kontrolle der Minimal- und Maximalwerte auf Plausibilität nach Angaben der jeweiligen Testmanuale, das korrekte Aufsummieren von Skalenwerten sowie die korrekte Berechnung von KL-Werten und Fehler-Prozent-Werten nach den Angaben in den jeweiligen Testmanualen.

Analyse der deskriptiven Statistiken. Der Anteil fehlender Werte, Verteilungsformen sowie Mittelwerte bzw. Mediane und Standardabweichungen aller Variablen wurden in Augenschein genommen. Um einzelne Variablen auf ihre Verteilungsformen zu prüfen wurden zum einen die nach West, Finch und Curran (1995) postulierten Grenzen von Schiefe < 2 und Kurtosis < 7 betrachtet. Zum anderen wurden für die Analysevariablen Histogramme mit zugehörigen Normalverteilungskurven generiert und geprüft. Mit den für nachfolgende Analysen ausgewählten Testvariablen wurden kreuzweise Streudiagramme

mit den zugehörigen Anpassungslinien der Funktion erstellt (jede Variable mit jeder). Diese Diagramme wurden dem Augenschein nach auf Ausreißerwerte überprüft.

Transformation der Daten. Um Punktwerte und Reaktionszeiten zusammen rechenbar zu machen, wurden die Daten z-transformiert. Jene Variablen, die Reaktionszeiten darstellen, wurden anschließend umgepolt (Multiplikation mit -1), sodass höhere Werte analog zu Punktwertevariablen bessere Testleistungen darstellten.

Überprüfung des Zusammenhangs mit Alter und Geschlecht. Die Zusammenhänge der Testvariablen mit den Variablen für Alter und Geschlecht wurden anhand von Korrelationen überprüft. Auch wenn die Zusammenhänge teilweise gering waren, wurden sie um den Anteil der beiden Variablen bereinigt. Hierfür wurden lineare multiple Regressionsanalysen mit Alter und Geschlecht als unabhängige Variablen für jeden einzelnen Analysewert als abhängige Variable verwendet und die standardisierten Residuen als weiter zu verwendender Variablenwert gespeichert. Die Überprüfung des Zusammenhangs der z-transformierten Variablen mit den standardisierten Residuen aus den linearen multiplen Regressionsanalysen wurde augenscheinlich mittels Streudiagrammen sowie rechnerisch über Korrelationen durchgeführt.

Schätzung fehlender Werte. Die Ergänzung fehlender Werte wurde mit der EM-Schätzung durchgeführt. Dieses Prozedere ist neben einem paarweisen oder listenweisen Fallausschluss die zu präferierende Methode (Lüdtke, Robitzsch, Trautwein & Köller, 2007). Der MCAR Test (= missing completely at random) nach Little (1988) testet zunächst, ob die Wahrscheinlichkeiten für fehlende Werte von den beobachteten Werten abhängen. Wird dieser nicht signifikant bedeutet dies, dass das Fehlen von Werten unabhängig ist von sowohl den fehlenden Werten als auch von allen anderen Werten im

Datensatz (Byrne, 2001). Der EM-Algorithmus setzt sich aus zwei Schritten zusammen. Der *Expectation*-Schritt vervollständigt fehlende Daten durch über die Regressionsgerade erwartete Werte. Im zweiten Schritt, *Maximization*, werden Modellparameter durch Maximierung der Likelihoodfunktion der vervollständigten Daten geschätzt. Diese beiden Schritte wechseln sich iterativ ab (Little & Rubin, 2002). Die Ergänzung fehlender Werte ist unabdingbar zur Durchführung des Bootstrapping Verfahrens in AMOS (Arbuckle, 2006). Dies ist wiederum erforderlich, wenn die multivariate Normalverteilung der Analysevariablen verletzt ist. Die Prüfung des Zusammenhanges der standardisierten Residuen, die fehlende Werte enthielten, mit den Variablen nach Ergänzung fehlender Werte wurde mittels Streudiagrammen in Augenschein genommen. Die Grafik gibt Aufschluss darüber, ob ein positiver linearer Zusammenhang vorliegt. Für die statistischen Analysen zur Schätzung von Reliabilitäts- und Validitätswerten wurden die standardisierten Residuen um fehlende Werte ergänzt weiter verwendet.

Prüfung der Testformen A und B sowie der Materialarten figural, numerisch und verbal des TEMEKKO auf Gleichheit via Strukturgleichungsmodelle. Aufgrund der nominellen Unterschiede der Testleistungen in den Formen A und B sowie in den unterschiedlichen Materialarten wurde geprüft, ob diese inhaltlich bedeutsam sind. Diese Überprüfung fand zunächst über Strukturgleichungsmodelle statt. Das Messmodell in Abbildung 3 diente der Inspektion beider Fragestellungen. Zum einen wurde die Gleichheit der beiden Testformen A und B und zum anderen die Gleichheit der drei Materialarten (figural, numerisch, verbal) geprüft. Charakteristisch für das Modell ist, dass die Fehlervarianzen sowie die Ladungen jeweils zu einander gleichgesetzt wurden. Diese Modellierung stellt ein essentiell-paralleles Modell dar. Von im wesentlichen (essentiellen) parallelen Messungen spricht man dann, wenn der wahre Wert in einem Test, Testteil oder Item für jede Person

lediglich um eine additive Konstante verschoben ist (Bühner, 2006, S. 125). Es wurde zur Prüfung der Tempowerte sowie der Fehler-Prozent-Werte angewendet.

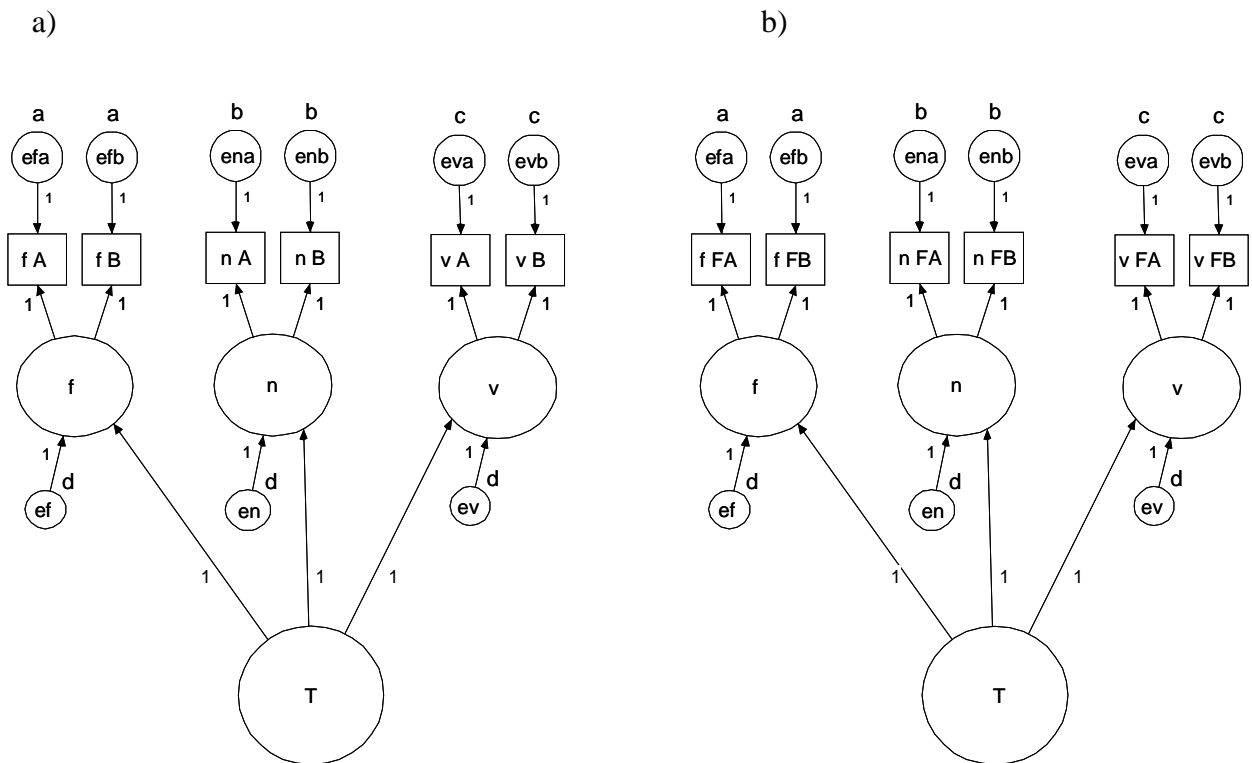


Abbildung 3: Messmodelle zur Prüfung der Testformen A und B sowie der Inhaltsdomänen. Modell a) Tempowerte, Modell b) Fehler-Prozent-Werte

Anmerkungen: a) T = TEMEKKO gesamt, f = figural, n = numerisch, v = verbal. fA = figural Form A, f B = figural Form B, n A = numerisch Form A, n B = numerisch Form B, v A = verbal Form A, v B = verbal Form B. e... = Fehlervariablen. b) T = TEMEKKO gesamt, f = figural, n = numerisch, v = verbal. f FA = figural Fehler-%-Wert Form A, f FB= figural Fehler-%-Wert Form B, n FA = numerisch Fehler-%-Wert Form A, n FB = numerisch Fehler-%-Wert Form B, v FA = verbal Fehler-%-Wert Form A, v FB = verbal Fehler-%-Wert Form B. e... = Fehlervariablen.

Aufgrund der Verletzung der multivariaten Normalverteilung (Prüfung mittels Mardia-Test in AMOS) in allen vier Validierungsstudien wurde der Modelltest unter Berücksichtigung der Bollen-Stine-Bootstrap-Methode durchgeführt. Die Verletzung der multivariaten Normalverteilung beeinträchtigt weniger die Schätzung der Parameter (Faktorladungen, Korrelationen). Vielmehr werden Pfade und Korrelationen zwischen manifesten Variablen, zwischen latenten Variablen oder zwischen manifesten und latenten Variablen aufgrund der zu geringen Schätzfehler signifikant, obwohl dies in der Population nicht zutrifft

(West, Finch & Curran, 1995). Dies kann dazu führen, dass passende Modelle zu oft abgelehnt werden. Um dieses Problem zu umgehen, wurde die Bollen-Stine-Bootstrap-Methode verwendet, die den p -Wert des χ^2 Tests korrigiert. Zur Beurteilung der Modellqualität wurden über den exakten Modelltest hinaus nachfolgende Fit Indizes gemäß der Empfehlung von Hu und Bentler (1999) verwendet: Standardized Root Mean Square Residual (*SRMR*) sowie Root Mean Squared Error of Approximation (*RMSEA*). Die von den Autoren angegebenen Cutoff-Werte dieser Indizes wurden berücksichtigt. So soll der *SRMR* $\leq .11$ sein, der *RMSEA* sollte $< .06$ sein für Stichproben mit mehr als 250 Personen und $< .08$ für $N < 250$. Darüber hinaus wurde der Comparative Fit Index (*CFI*) als Bewertungskriterium herangezogen, wie von Beauducel und Wittmann (2005) empfohlen. Dieser sollte $\geq .95$ sein.

Alternative Prüfung auf Gleichheit der Testformen A und B bzw. der Materialarten (figural, numerisch, verbal) via Kovarianzanalysen. Um zu prüfen, ob nominell vorhandene Unterschiede der verwendeten Materialien und Testformen A und B bedeutsam sind, wurden alternativ die Mittelwerte auf deren Gleichheit zueinander untersucht. Für die Prüfung der Materialarten zueinander wurden drei Kovarianzanalysen jeweils für einen paarweisen Vergleich berechnet, wobei für die Variablen *Alter* und *Geschlecht* kontrolliert wurde (figural : numerisch, numerisch : verbal, figural : verbal). Darüber hinaus wurde ein zusätzlicher Vergleich aller drei Materialarten angestellt. Die zu prüfende Nullhypothese lautete: es liegen keine Mittelwertsdifferenzen vor. Aus statistischer Sicht sollte daher als strenges Maß die Alternativhypothese (die Materialarten sind unterschiedlich) so früh wie möglich angenommen werden können. Zur Beurteilung der empirischen Effektgröße wurde für jeden einzelnen Vergleich der kritische F -Wert für ein partielles η^2 von .01 (kleinstmöglich angenommener Effekt) berechnet. Darüber hinaus wurden die angenommenen Teststärken von 95 % bei den paarweisen

Einzelvergleichen aufgrund der Mehrfachvergleiche der Anpassung nach Bonferroni-Holm (Holm, 1979) unterzogen, um einer Alpha-Inflation entgegenzuwirken. Dieser Ansatz gilt vor allem bei großen Stichprobengrößen als sehr konservativ in Bezug auf das Verwerfen der Nullhypothese. Bei den paarweisen Vergleichen wurde die Teststärke von 95 % entsprechend auf $.983 (1 - \beta/3)$ für den stärksten empirischen Effekt, auf $.975 (1 - \beta/2)$ für den zweitstärksten empirischen Effekt und auf $.95 (1 - \beta)$ für den dritten Effekt festgelegt; wobei β aufgrund der Prüfung auf Gleichheit (nicht auf Verschiedenheit) auf $.05$ festgesetzt wurde. Die Teststärke $(1 - \beta)$ für den 3-fach Vergleich wurde auf 95 % determiniert. Der kritische F -Wert für die kleinstmöglich angenommene Effektgröße ($\eta^2 = .01$) wurde mit dem empirischen F -Wert verglichen. Lag der empirische Wert außerhalb der kritischen Grenze, so musste die Nullhypothese (Gleichheit) abgelehnt werden. In solchen Fällen wurden die entsprechenden Effektstärken einer detaillierteren Interpretation unterzogen. Als Grenzen für das Effektstärkemaß η^2 galten die Konventionen nach Cohen (1988), i.e. $\eta^2 \leq .01$ kleiner Effekt, $\eta^2 \leq .059$ mittlerer Effekt sowie $\eta^2 \leq .138$ großer Effekt. Analog zu diesem Vorgehen verhielt es sich mit der Prüfung auf Gleichheit der Testformen A und B auf dem Wege der Kovarianzanalysen. Naturgemäß wurden hier lediglich drei Vergleiche angestellt; i.e. figural A : figural B; numerisch A : numerisch B sowie verbal A : verbal B.

Reliabilitätsschätzungen. Für die Auswahl des geeigneten Reliabilitätsmaßes waren verschiedene Überlegungen zu berücksichtigen. Einerseits gilt die Retestreliabilität als geeignete Schätzung, wenn es sich um Schnelligkeits-Tests (sog. Speed-Tests) handelt bzw. bei Persönlichkeits- und Leistungstests, die zeitlich stabile Merkmale erfassen sollen (Bühner, 2006, S. 128). Dem gegenüber aber ist die Berechnung einer Retestreliabilität nur sinnvoll, wenn Lern- oder Übungseinflüsse etc. weitgehend ausgeschlossen werden können, die Durchführungsbedingungen zu beiden Testzeitpunkten gleich sind und ein

nach praktischen Gesichtspunkten angemessenes Retestintervall gewählt werden kann. Andererseits ist auch die Schätzung der Paralleltestreliabilität ein geeignetes Maß bei Speed-Tests (Bühner, 2004, S. 141). Die zugrunde liegende Annahme (vgl. 2.1 Konzentration), Konzentration als einen Zustand zu erfassen, dessen Messung mit situativen Faktoren sowie mit Personenmerkmalen kovariiert, gleichwohl die Fähigkeit aber als Persönlichkeitsmerkmal aufgefasst wird, spricht nicht eindeutig für die Notwendigkeit der Retestrelia­bilität. Darüber hinaus berichtet Bühner (2001) von *unübersehbaren Problemen* bei der Retestung mit Konzentrationstests, aufgrund von Übungs- und Transfereffekten. Wenngleich solche Effekte durch die besondere Konstruktionsweise der Testitems des TEMEKKO nicht zu erwarten sind, wurden sie bislang nicht untersucht und können demnach nicht zuverlässig ausgeschlossen werden. Da die beiden Testformen A und B des TEMEKKO essentiell-parallele Formen darstellten und die Anwendung der Paralleltestkorrelation auch bei Speed-Tests als geeignetes Maß gilt, wurde als Maß zur Schätzung der Reliabilität für kleine Stichprobengrößen die Formel nach Kristof (1963) verwendet, die von der Annahme eines parallelen Modells ausgeht. Es wurden die Reliabilitätsschätzungen der Tempowerte und Fehler-Prozent-Werte berechnet.

$$r_{tt} = \frac{2}{N - 1} + \frac{N - 3}{N - 1} \cdot \left[\frac{4 \cdot S_1 \cdot S_2 \cdot r_{12}}{S_1^2 + S_2^2 + 2 \cdot S_1 \cdot S_2 \cdot r_{12}} \right]$$

Abbildung 4: Formel von Kristof (1963) zit. nach Bühner (2006)

Anmerkungen: r_{tt} = Reliabilität, N = Anzahl der Probanden, r_{12} = Korrelation zwischen Testform A und Testform B, S_1 = Standardabweichung der Testform A, S_2 = Standardabweichung der Testform B.

Validitätsschätzungen. Zunächst wurden die Korrelationsmatrizen der einzelnen Studien in Augenschein genommen, was erste Aufschlüsse über Zusammenhänge mit bzw. Abgrenzungen zu den Testleistungen des TEMEKKO erkennen ließ.

Zur Vorhersage der TEMEKKO Testleistung wurden lineare multiple Regressionsanalysen (Einschluss Methode) sowie latente Regressionsanalysen berechnet. Die in den einzelnen Studien vorhandenen Kennwerte anderer Verfahren flossen als Prädiktoren ein. Die TEMEKKO Testleistung diente als Kriterium. In manchen Studien wurden die Regressionsanalysen auf Parcel-Ebene durchgeführt. Die Vorteile der Aggregatsbildung liegen zum einen darin, dass ein Aggregat eine höhere Reliabilität aufweist, als einzelne Indikatoren. Zum anderen hat ein Parcel einen höheren Prozentsatz an gemeinsamer Varianz im Vergleich zu spezifischer Varianz. Darüber hinaus besitzen Aggregate meist günstigere Verteilungen als einzelne Variablen (Little, Cunningham, Shahar & Widaman, 2002, S. 154). Zur Schätzung der konvergenten und divergenten Validität bzw. zur Verortung der TEMEKKO Testleistung in einer Struktur der kognitiven Leistungsmaße der Intelligenz, Aufmerksamkeit und der Arbeitsgedächtniskapazität wurden Messmodelle in Anlehnung an in der Forschung bestehende und etablierte Modelle sowie aus theoretischen Überlegungen zu den analysierten Konstrukten generiert und mittels konfirmatorischer Faktorenanalysen auf Passung der Daten geprüft.

Beispielhaft sei das Messmodell zur Aufmerksamkeitsstruktur nach Moosbrugger, Goldhammer & Schweizer (2006) auch hinsichtlich seines theoretischen Hintergrundes ausführlich beschrieben, das die Grundlage der Analyse in Studie 1 bildete. Dieses Modell ist zum einen das bislang einzige Modell, das aufgrund einer systematischen Überprüfung verschiedener Aufmerksamkeits-tests und Modelle entstanden ist. Zum anderen wurde es anhand von gesunden Probanden entwickelt. Das Modell der Aufmerksamkeit nimmt eine globale Unterscheidung der Komponenten *Perceptual Attention* und *Executive Attention* vor. Der erstgenannte Faktor reflektiert das Erkennen und die Selektion von Reizen. Der Faktor *Executive Attention* hingegen spiegelt die intentionale Kontrolle innerer Prozesse wider, die an der Selektion beteiligt sind. Das Messmodell ist so strukturiert, dass alle

verwendeten Analysevariablen auf dem Faktor Perceptual Attention laden, der somit als Generalfaktor aufgefasst werden kann und die gemeinsame Varianz aller verwendeten Maße abbildet. Testvariablen, von denen angenommen wird, dass sie exekutive Kontrolle erfordern, sind zusätzlich Indikatoren des Faktors Executive Attention. Der Ordnung halber sei nicht unerwähnt, dass Moosbrugger et al. (2006) in ihrer Untersuchung ausschließlich Reaktionszeitmaße verwendet haben, sodass sie den beiden Komponenten genauer gesagt die Geschwindigkeit (speed) von Perceptual bzw. Executive Attention zuschreiben. Die exekutive Kontrolle der Theory of Visual Attention (TVA) (Bundesen, 1990) stellt das Entscheidungskriterium dar, ob ein Verfahren dem einen oder dem anderen latenten Faktor zugeordnet werden soll. Die TVA nimmt an, dass beides, visuelle Erkennung und Selektion von Objekten auf wahrnehmungsgebundener Kategorisierung basieren, d.h. dass das Objekt einer von mehreren Kategoriealternativen zugeordnet wird. Die hierfür benötigte Zeit hängt von mehreren Parametern ab (Stärke des sensorischen Eindrucks einer Kategorie, Stärke der Aufmerksamkeit eines Objekts und subjektive Neigung, ein Objekt einer bestimmten Kategorie zuzuordnen (decision bias)). Der exekutive Prozess, der diese untergeordnete Stimulusselektion gemäß der TVA steuert, ist in der Theorie von Logan & Gordon (2001) dargestellt. Diese Theorie stellt darüber hinaus auch den untergeordneten Prozess der Response-Selection (Antwortauswahl) dar, wie vom Exemplar-Based-Random-Walk (EBRW) Modell (Nosofsky & Palmeri, 1997) postuliert. Die TVA bezieht sich auf die Selektion innerhalb einer Aufgabe, wohingegen die exekutive Kontrolle der TVA (EC-TVA) die Selektion zwischen den Aufgaben erklärt. Diese Aufgaben-Sets sind im Arbeitsgedächtnis und in untergeordneten TVA Prozessen vertreten. Um ein Aufgabenset (z.B. decision bias) zu bilden, müssen die Kontrollparameter vom Arbeitsgedächtnis zu untergeordneten Prozessen (TVA oder EBRW) transferiert werden. Um es zu verändern, muss ein neues Set vom Arbeitsgedächtnis zu untergeordneten Prozessen übertragen werden. Aus dieser

theoretischen Annahme lassen sich Aufmerksamkeitstests nach ihrem Fokus einer der beiden Komponenten zuordnen. Für Perceptual Attention ist der Wahrnehmungsaspekt (Posner, 1980; Posner & Rafal, 1987; Bundesen, 1990) der Aufmerksamkeit maßgeblich (z.B. einfache Signalentdeckung gemäß TVA), für Executive Attention hingegen ist die exekutive Kontrolle (Norman & Shallice, 1986; Baddeley, 1986; Engle, Tuholski, Laughlin & Conway, 1999) als Aspekt der Aufmerksamkeit entscheidend (Steuerung untergeordneter Prozesse gemäß EC-TVA).

Das nun anschließende Kapitel befasst sich mit der Darstellung der Hauptgütekriterien. Zur Ergänzung der Berichte über die Reliabilitätsschätzungen werden zu Beginn dieses Unterpunktes die geprüften Voraussetzungen für das gewählte Reliabilitätsmaß dargelegt. Die Ausführungen zur Validitätsschätzungen des TEMEKKO werden gespeist von den Analysen aus den zu Grunde liegenden Validierungsstudien 1 bis 4 und stellen eine Zusammenfassung der Ergebnisse der Befunde dar. Für vertiefende Einblicke in die Berechnungen sei auf die Anhänge I bis IV verwiesen.

5 Hauptgütekriterien

5.1 Objektivität

Ein hohes Maß an Durchführungsobjektivität ist durch die Standardisierung der Untersuchungssituation und des präzise festgelegten Ablaufs der Testung gegeben (vgl. 3.3 Durchführung) sowie die verbindlichen standardisierten Instruktionen des TEMEKKO. Durch festgelegte und bindende Auswertungsvorschriften sind alle Items entweder eindeutig richtig oder eindeutig falsch gelöst. Die Ermittlung der Bearbeitungsmenge und KL-Werte über Schablonen und Berechnung der Fehler-Prozent-Werte über festgelegte Formeln macht die Auswertung unabhängig vom Auswerter. Somit ist die maximale Auswertungsobjektivität gewährleistet (vgl. 3.4 Auswertung). Darüber hinaus erhöht die Möglichkeit einer Gruppentestung die Vergleichbarkeit der Testergebnisse, sofern externe, nicht zu kontrollierende Einflüsse wirksam würden. Die Interpretationsobjektivität ist aufgrund der für den Gesamttest und alle Testteile vorhandenen Normtabellen (vgl. Anhang VI) sichergestellt.

5.2 Reliabilität

Zunächst seien die Befunde der Strukturgleichungsanalysen als Belege dafür benannt, dass es sich bei den Testformen A und B des TEMEKKO um Parallelförmigkeiten handelt. Darüber hinaus konnte belegt werden, dass die Testleistungen der Untertests vom verwendeten Material relativ unabhängig sind. Die für alle Validierungsstudien verwendeten Messmodelle (4.1.2 Prozedere der Datenbearbeitung) zeigten in allen vier Untersuchungen

einen exakten Modellfit, sowohl für die Tempowerte als auch für die Fehler-Prozent-Werte (vgl. Anhänge I, II, III und IV jeweils 4.4 Prüfung auf Gleichheit der Testformen und Materialarten). Die Testformen A und B können demnach als essentiell-parallele Formen anerkannt werden. Außerdem lässt sich festhalten, dass auch die verwendeten Materialarten (verbal, numerisch und figural) der Untertests das gleiche Konzept - gleich gut - abbilden. Leistungsunterschiede in den Subtests lassen sich nicht auf eine bestimmte Materialart zurückführen. Tabelle 4 zeigt einen Überblick über die Fit-Indizes für beide Messmodelle; für Tempowerte sowie für Fehler-Prozent-Werte.

Tabelle 4: Ergebnisse der Modellprüfungen auf Gleichheit der Testformen und Materialarten

Tempowerte	χ^2 [16]	<i>p</i>	<i>CFI</i>	<i>RMSEA</i>	<i>SRMR</i>
Studie 1	27.29	.13	.98	.07 [.02; .12]	.07
Studie 2	22.40	.30	.99	.05 [.00; .10]	.03
Studie 3	28.10	.17	.98	.08 [.02; .12]	.05
Studie 4	18.93	.31	.99	.04 [.00; .10]	.07

Fehler-%-Werte	χ^2 [16]	<i>p</i>	<i>CFI</i>	<i>RMSEA</i>	<i>SRMR</i>
Studie 1	20.71	.66	.98	.05 [.00; .10]	.06
Studie 2	11.62	.94	1.00	.00 [.00; .05]	.04
Studie 3	10.95	.95	1.00	.00 [.00; .05]	.04
Studie 4	16.57	.73	1.00	.02 [.00; .09]	.05

Alternative Prüfungen der Mittelwertsdifferenzen der Tempowerte via Kovarianzanalysen für die Variablen *Alter* und *Geschlecht* kontrolliert, replizierten das Bild im Wesentlichen. Unter den strengen Kriterien der Prüfung auf Gleichheit (nicht auf Unterschiedlichkeit), unter der Annahme einer Teststärke von 95 % sowie der Anpassung derselben nach Bonferroni-Holm zur Abwendung einer Alpha-Inflation aufgrund von Mehrfachvergleichen, zeigten sich einige wenige Unterschiede. Von den bedeutsamen Unterschieden über alle untersuchten Stichproben hinweg wies der überwiegende Teil

lediglich geringe Effektstärken auf, i.e. $\eta^2 = .02$ für die Materialvergleiche numerisch : verbal und figural : verbal in Studie 2, $\eta^2 = .02$ für den Materialvergleich figural : verbal sowie $\eta^2 = .03$ für den 3-fach Vergleich figural : numerisch : verbal in Studie 3 sowie mit $\eta^2 = .02$ für die Materialvergleiche figural : numerisch und numerisch : verbal und mit $\eta^2 = .03$ für den 3-fach Vergleich figural : numerisch : verbal in Studie 4; außerdem mit $\eta^2 = .03$ in Studie 4 für die Vergleiche der Testformen A und B in den figuralen und verbalen Bereichen. Mittlere Effektgrößen erreichten die Vergleiche der Testformen A und B des figuralen Materials in Studie 1 ($\eta^2 = .07$) sowie die Materialvergleiche figural : numerisch in Studie 3 ($\eta^2 = .05$) und der Materialvergleich figural : verbal in Studie 4 ($\eta^2 = .08$). Alle übrigen Vergleiche über alle verwendeten Studien hinweg zeigten weder statistische Signifikanz noch praktisch bedeutsame Unterschiede. Eine detaillierte Beschreibung der Befunde findet sich in den Anlagen I bis IV (4.4 Prüfung auf Gleichheit der Testformen und Materialarten).

Wie in Abschnitt 4.1.2 (Prozedere der Datenbearbeitung) erläutert, wurde die Paralleltestkorrelation nach Kristof (1963) als geeigneter Schätzer der Reliabilität gewählt. Nachfolgende Tabelle zeigt die Reliabilitätsschätzungen der Tempo-Werte (KL-Werte) sowie der Fehler-Prozent-Werte über alle vier Validierungsstudien hinweg.

Tabelle 5: Paralleltestreliabilität (unverzerrte Schätzer)

	Studie 1 (<i>N</i> = 138)	Studie 2 (<i>N</i> = 144)	Studie 3 (<i>N</i> = 128)	Studie 4 (<i>N</i> = 121)
figural	.79	.90	.89	.85
numerisch	.89	.96	.94	.92
verbal	.90	.92	.91	.89
gesamt	.93	.97	.96	.94
Auslassungsfehler %	.82	.89	.89	.89
Verwechslungsfehler %	.89	.70	.64	.51
Fehler gesamt %	.88	.88	.88	.88

Die Höhe der Reliabilitätsschätzungen kann mit Werten von $r_{tt} = .93$ bis $r_{tt} = .97$ für den Gesamttest (KL-Wert gesamt) als sehr gut eingestuft werden (Fisseni, 1997, S. 124). Die Reliabilitätswerte für die Untertests liegen mit $r_{tt} = .89$ bis $r_{tt} = .96$ für das numerische und mit $r_{tt} = .89$ bis $r_{tt} = .92$ für das verbale Material im (sehr) hohen Bereich. Lediglich der Wert der Studie 1 für den figuralen Subtest liegt mit $r_{tt} = .79$ unter den Erwartungen. Diese Stichprobe setzte sich aus Personen im fortgeschrittenen Alter ($M = 69.5$ Jahre) zusammen. Die übrigen Studien zeigen im figuralen Bereich mit $r_{tt} = .85$ bis $r_{tt} = .90$ Reliabilitätsschätzungen, die als gut bis sehr gut bezeichnet werden können.

Die Reliabilitäten der Fehler-Prozent-Werte lagen für alle Studien in überraschend hohen Bereichen. Erwartungskonform wiesen die Verwechslungsfehler-Prozent-Werte ($r_{tt} = .51$ bis $r_{tt} = .89$) niedrigere Reliabilitäten auf als jene der Auslassungsfehler-Prozent-Werte ($r_{tt} = .82$ bis $r_{tt} = .89$). Die Fehler-Prozent-Werte gesamt erreichten Reliabilitäten von $r_{tt} = .88$. Diese Reliabilitätsschätzungen lassen sich als Maß für die Stabilität begreifen, mit der auch Fehler gemacht werden. Dies spricht für die Homogenität der Testitems, gleichzeitig aber auch für die homogen verteilten erschwerenden Bedingungen (Zusatzmerkmale der Reize) bei der Bearbeitung.

5.3 Validität

Die *Inhaltsvalidität* eines Verfahrens lässt sich statistisch nicht absichern, sondern lediglich aus theoretischer Sicht bestimmen. Ein Test gilt dann als inhaltlich valide, wenn seine Elemente so beschaffen sind, dass sie den zu erfassenden Messgegenstand repräsentieren (Lienert & Raatz, 1998). Der Messgegenstand wird durch die dem Verfahren zu Grunde liegende Definition von Konzentration nach Schmidt-Atzert,

Büttner & Bühner (2004) sowie Schmidt-Atzert, Krumm & Bühner (2008) bestimmt (vgl. 2.1 Konzentration). Die Testaufgaben des TEMEKKO beinhalten alle in der Definition enthaltenen Merkmale und operationalisieren die in Abschnitt 2.2 (Konzentrationstests Abs. 2) genannten Anforderungen an Probanden. Vor dem Hintergrund der Einschränkung auf die Durchführung als Durchstreichetest können die Items als repräsentative Stichprobe eines theoretisch möglichen Itempools in o.g. Sinne angesehen und eine Inhaltsvalidität angenommen werden.

Erste Untersuchungen zur Schätzung der *Konstruktvalidität* des TEMEKKO als Verfahren zur Erfassung von Konzentration bzw. exekutiver Kontrolle umfassten vier voneinander unabhängige Studien. Nachfolgend seien die Ergebnisse der Analysen nach inhaltlichen Gesichtspunkten zusammengefasst.

Indizien für die Konvergenz des TEMEKKO mit Konzentration bzw. exekutiver Kontrolle.

Hierfür gab es in jeder der 4 Studien beachtliche Befunde. Der TEMEKKO markierte den Faktor Executive Attention (Studie 1) mit $\lambda = .79$, gefolgt vom Test d2 ($\lambda = .60$) und Shifting (Reaktionswechsel, TAP) ($\lambda = .49$). Die Untertests des TEMEKKO erwiesen sich auf funktionaler Ebene als die stärksten Indikatoren für exekutive Kontrolle (Studie 2) mit ($\lambda = .89$ numerisch, $\lambda = .79$ verbal und $\lambda = .75$ figural). Auf latenter Ebene (Studie 3) betrug der gemeinsame Varianzanteil des TEMEKKO mit Konzentration 74 % (bzw. 81 %, wenn der Einfluss von Materialvarianzanteilen unberücksichtigt blieb). Die TEMEKKO Testleistung konnte rechnerisch nicht von den Konzentrationstests (d2 und FAKT) separiert werden (Studie 4) (minderungskorrigierte Korrelation $r \sim 1$).

Zusammenhänge zwischen TEMEKKO und anderen Konstrukten. Besonders enge Zusammenhänge erwiesen sich in Studie 2 auf funktionaler Ebene durch die hohe

Korrelation ($r = .80$) des Faktors exekutive Kontrolle mit Relational Integration. Studien 3 und 4 wiesen Korrelationen des TEMEKKO zu Intelligenz von $r = .77$ bzw. $r = .76$ auf. Die breiter konzipierte Analyse in Studie 4 zeigte darüber hinaus hohe Zusammenhänge mit Coordination ($r = .68$), wie auch zwischen den Konstrukten ($r = .74$ zwischen Intelligenz und Coordination; $r = .69$ zwischen Intelligenz und Speichern im Kontext gleichzeitiger Verarbeitung). Noch deutlicher wurde der Zusammenhang zwischen Intelligenz, Coordination und dem TEMEKKO nach Berücksichtigung eines übergeordneten Faktors, der aufgrund der besonders hohen Ladung von Intelligenz ($\lambda = .98$) als allgemeine Intelligenz verstanden wurde. Dann nämlich waren der TEMEKKO ($\lambda = .79$) und Coordination ($\lambda = .78$) gleichermaßen dort zu verorten.

Prädiktion der TEMEKKO Testleistung. Die Vorhersage des erklärbaren Varianzanteils des TEMEKKO wurde in erster Linie von Maßen der Konzentration und Intelligenz bestimmt. Darüber hinaus kam auch der Umstellfähigkeit (shifting) ein gewisses Maß an Bedeutung zu. Die nachstehende Tabelle 6 zeigt die Vorhersagewerte im Überblick.

Tabelle 6: Vorhersagen der TEMEKKO Testleistung über 4 Studien

	erklärbarer Varianzanteil	Prädiktor	Konstrukt	einzigartiger Anteil
Studie 1	40 %	Test d2	Konzentration	14 %
		Reaktionswechsel	Shifting	3 %
		Geteilte Aufmerksamkeit	selektive Aufmerksamkeit	3 %

Tabelle 6 wird fortgesetzt

5 Hauptgütekriterien

Studie 2	46 %	Parcel (I-S-T 2000 R)	Intelligenz	11 %
		Parcel (Switching Numerisch, Geteilte Aufmerksamkeit, Go/Nogo)	kognitive Kontrolle	5 %
Studie 2	24 %	Switching Numerisch	Shifting	16 %
Studie 3	54 %	Parcel (Konzentrationstests)	Konzentration	15 %
		Parcel (Intelligenztests)	Intelligenz	7 %
Studie 4	65 %	d2 / FAKT	Konzentration	21 %
		Parcel (I-S-T 2000 R)	Intelligenz	6 %
Studie 4	66 %	Parcel (I-S-T 2000 R)	Intelligenz	$\beta = .69$
		Switching Kosten	Supervision/ Shifting	$\beta = .22$

Prädiktion der Intelligenz. Aus Perspektive der Intelligenz als Kriterium konnten 9 % ihres erklärbaren Varianzanteils (38 %) auf den TEMEKKO zurückgeführt werden (Studie 3). In Studie 4 lag auf latenter Ebene seine spezifische Vorhersage bei $\beta = .42$ bei einem erklärbaren Anteil von 79 % der abhängigen Variable. Damit konnte der TEMEKKO ebenso gut Intelligenz vorhersagen wie das Konzept Speichern im Kontext der Verarbeitung ($\beta = .42$).

Abhängigkeit von Inhaltsdomänen. Die Modellierung in Studie 2 machte die Kontrolle für figurales Material unter Berücksichtigung des entsprechenden TEMEKKO Untertests erforderlich ($r = .31$). Auch in Studie 4 wies der figurale Untertest des TEMEKKO eine maßgebliche Ladung auf dem Materialfaktor auf ($\lambda = .42$). In den übrigen Studien sowie für die übrigen Materialarten wurden keine nennenswerten Zusammenhänge festgestellt.

6 Nebengütekriterien

Es folgen nun Ausführungen zu den Nebengütekriterien. Die Darstellung der Charakteristika der Stichproben zur Ermittlung der Normwerte sowie die Berechnung derselben bilden den ersten Teil dieses Kapitels. Erläuterungen zur Ökonomie sowie zur Akzeptanz des TEMEKKO schließen diesen Punkt.

6.1 Normierung

Für den TEMEKKO wurden Normtabellen (Anhang VI) für die KL-Werte getrennt nach Testformen A und B und nach Inhaltsdomänen sowie für die KL-Werte der Formen A und B und den Gesamttest bezogen auf zwei von einander unabhängigen Stichproben (Schüler und Senioren) zur Verfügung gestellt. Darüber hinaus liegen die Standardwerte für Fehlerprozentage für die Testformen A und B und für den Gesamttest für die jeweilige Stichprobe vor.

Die Stichprobe der Schüler umfasste 312 Personen im Alter von 16 bis 45 Jahren ($M = 21.8$ Jahre; $SD = 3.0$ Jahre). Das mittlere Alter der 127 Frauen (40.7 %) betrug 21.8 Jahre ($SD = 2.1$ Jahre), das der 185 Männer ebenso 21.8 Jahre ($SD = 3.4$ Jahre). Die Daten wurden im Mai 2007 in Berliner Schulen erhoben. Lediglich 148 Personen (47.4 %) machten nähere Angaben zum Schultyp, den sie besuchten. Demnach waren 143 Berufsschüler, 3 Fachschüler sowie je eine Person der Realschule bzw. Fachoberschule zuzuordnen. Diese Stichprobe kann dem Augenschein nach hinsichtlich Alter und Bildungsstand als homogen angenommen werden.

Für die Altersgruppe > 60 Jahre diente die Stichprobe der Validierungsstudie 1 als Basis vorläufiger Normen. Sie umfasste 138 Personen im Alter von 59 bis 92 Jahren ($M = 69.5$ Jahre; $SD = 7.17$ Jahre). Das mittlere Alter der 89 weiblichen Versuchspersonen (64.5 %) betrug 70.1 Jahre ($SD = 7.59$ Jahre), das der 49 männlichen Probanden 68.6. Jahre ($SD = 6.29$ Jahre). Für die Erstellung der Normen der KL-Werte dienten die Daten von 136 Personen und für die Fehler-Prozent-Werte Normen von 135 Personen dieser Stichprobe, die übrigen 2 bzw. 3 Fälle hatten fehlende Werte.

Die Berechnung der Standardwerte erfolgte nach folgendem Schema. Zunächst wurden die KL-Werte anhand der Mittelwerte und Standardabweichungen der zugrunde gelegten Normierungsstichprobe (Schüler bzw. Senioren) nach folgender Formel in z-standardisierte Werte transformiert.

$$z = (KL - M) / SD$$

Abbildung 5: Formel zur Berechnung des z-Wertes

Anmerkungen: z = z-standardisierter Wert, KL = KL-Wert, M = Mittelwert der Normierungsstichprobe, SD = Standardabweichung der Normierungsstichprobe

Im nächsten Schritt wurden die z-transformierten Werte in Standardwerte ($M = 100$, $SD = 10$) nach folgender Formel umgewandelt.

$$Z = z * 10 + 100$$

Abbildung 6: Formel zur Berechnung des Standardwertes

Anmerkungen: Z = Standardwert, z = z-standardisierter Wert

Die Ermittlung der Standardwerte der Fehler-Prozente wurde analog vorgenommen. Die errechneten Standardgrößen wurden abgerundet, bei $< .5$ und aufgerundet bei $\geq .5$.

6.2 Ökonomie

Durch die Verwendung eines gebundenen Antwortformats (Durchstreichen der Aufgaben) kann die Bearbeitungs-, Auswertungs- und Lösungszeit kurz gehalten werden. Die bereitgestellten Schablonen zur Auswertung tragen ebenso zur Ökonomie des Verfahrens bei wie die Möglichkeit einer Gruppentestung.

6.3 Akzeptanz

Die Befragung einiger zufällig ausgewählter Probanden, deren Leistungen allerdings nicht in die statistischen Analysen eingingen, ergab generell eine große Aufgeschlossenheit gegenüber der Testung. Nichtsdestotrotz wurde berichtet, dass die Bearbeitung des Verfahrens als sehr anstrengend erlebt wurde und viel Konzentration abverlangte. Die Schwierigkeit der Reizeverarbeitung werde aufgrund der prinzipiell trivialen Aufgabenstellung unterschätzt, würde aber insbesondere durch die geforderte Aufrechterhaltung der Fokussierung und die Ablenkung durch die Zusatzmerkmale deutlich. Systematische Untersuchungen zur Akzeptanz des TEMEKKO anhand von geeigneten Methoden stehen gegenwärtig noch aus.

Im Allgemeinen kann die Akzeptanz der Testung bei Probanden durch eine sachgemäße Anwendung und Durchführung sowie transparente Rückmeldung der Testleistungen erhöht bzw. vorausgesetzt werden.

Das nun folgende Fazit stellt eine kurze Ableitung als Schlussbetrachtung dar.

7 Fazit

Die Befunde zeigen auf eindrucksvolle Weise, dass es gelungen ist, theoriegeleitet ein Verfahren zu entwickeln, das die gängigsten Merkmalskriterien von Konzentration operationalisiert und das auf einer konkreten Definition des Konstrukts basiert. Der TEMEKKO erfasst Konzentration auf Konstruktebene bzw. exekutive Kontrolle aus funktionaler Sicht.

Die Anforderung, die der TEMEKKO an Probanden stellt, ist zunächst eine basale wahrnehmungsgebundene Aufmerksamkeitsleistung, die als notwendige, aber nicht hinreichende Grundlage der höheren kognitiven Leistung gilt. Darauf aufbauend erfordert der TEMEKKO Fähigkeiten, die eine konzentrierte Leistung und Überwachung fortlaufender kognitiver Prozesse und Handlungen, die intentionale Kontrolle innerer Prozesse bei Reizselektion sowie die Unterdrückung von irrelevanten Stimuli darstellen.

Aufgrund der sehr guten Reliabilitäten des TEMEKKO sowie der belegten Konstruktkonvergenz mit Konzentration kann das Verfahren als Markertest zur Erfassung von Konzentration bzw. exekutiver Kontrolle bezeichnet werden.

Die Bereitstellung von Normen für junge Erwachsene und für Senioren ermöglicht beim Einsatz des TEMEKKO in der Praxis die Einbettung einer individuell erhobenen Fähigkeitsausprägung in einen Bezugsrahmen. Für die Forschung eignet sich der TEMEKKO aufgrund seiner fundierten Konstruktionsweise und Validierung in besonderem Maße.

8 Perspektiven

Weitere Untersuchungen zur Validität wären wünschenswert. Zum besseren Verständnis der funktionellen Überlappung bzw. Abgrenzung zum Konstrukt Intelligenz wäre die Verwendung des APM Advanced Progressive Matrices Tests (Raven, 1976), der häufig als Goldstandard der Intelligenztests bezeichnet wird, angezeigt. Aufgrund der offenkundigen Anforderung des TEMEKKO an Inhibition wäre eine Betrachtung zur Abgrenzung zum FWIT Farb-Wort-Interferenztest (Bäumler, 1985) sinnvoll, bzw. zum - zumindest für den numerischen Teil konzeptionell nahezu kongruenten - numerischen Zahlenstroop-Test. Zweckmäßig wäre auch eine Untersuchung zum Einfluss motivationaler Aspekte auf die TEMEKKO Testleistung als konzentrierte Leistung. Die Durchführung von Tests zur sensorischen Leistungsfähigkeit oder zu motorischen Voraussetzungen als Vortests zum TEMEKKO können - nach eingehender Prüfung - angedacht werden.

Darüber hinaus fehlen bislang Untersuchungen in praktischen Einsatzgebieten. Die im Überblick dieser Arbeit benannten Anwendungsbereiche begründen sich auf theoretischen Überlegungen zum Konzept Konzentration und zu exekutiven Funktionen. Insbesondere fehlen Untersuchungen in der Eignungsdiagnostik mit Befunden zu prädiktiver Validität im Hinblick auf Berufserfolg. Im klinischen Bereich liegen weder Untersuchungsergebnisse noch Normdaten vor. Dabei kann spekuliert werden, dass sich der TEMEKKO aufgrund seiner einfachen Aufgaben und relativ kurzen Bearbeitungszeit pro Testform gut zum Einsatz bei Patienten mit z.B. kognitiver Leistungsminderung (im Alter) oder Erkrankungen, wie affektiven Störungen oder Schizophrenie und zugehöriger Störungen eignet.

Die angedachte Verwendung des TEMEKKO im Bereich der pädagogischen Psychologie basiert ebenso lediglich auf theoretischen Überlegungen. Untersuchungen, insbesondere vor dem Hintergrund von noch nicht vollständig entwickelten exekutiven Funktionen in der Kindheit und Jugend, wären in diesem Zusammenhang wünschenswert.

Es sei darauf hingewiesen, dass die vorliegenden Normen lediglich vorläufigen Charakter besitzen. Weitere Untersuchungen auf einer breiteren Basis i.S.v. breiter gestreuten demografischen Merkmalen wären angezeigt. Insbesondere für Kinder liegen bislang keine Anhaltspunkte vor. Daher ist das Verfahren bislang nur für die Anwendung bei Erwachsenen vorgesehen.

9 Literaturverzeichnis

Abels, D. (1974). *Konzentrations-Verlaufs-Test (KVT)*. (2. Aufl.). Göttingen: Hogrefe.

Amelang, M. (1996). Intelligenz. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie der Bereiche Differentielle Psychologie und Persönlichkeitsforschung, Band 2: Verhaltens- und Leistungsunterschiede* (S. 245 – 328). Göttingen: Hogrefe.

Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R) - Handanweisung*. Göttingen: Hogrefe.

Andreasen, N. (2002). *Brave New Brain. Geist Gehirn Genom*. Berlin: Springer.

Arbuckle, J. L. (2006). Amos (Version 7.0) [Computer Program]. Chicago: SPSS.

Baddeley, A.D. & Hitch, G. (1974). Working memory. In g.H. Bower (Ed.), *The Psychology of Learning and Motivation*, 47 – 90. San Diego: Academic Press.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford Univ. Press.

Baddeley, A., & Della Salla, S. (2003). Working memory and executive control. In: Roberts A. C., Robbins, T. W., Weiskrantz, L. (Hrsg). *The prefrontal cortex. Executive and cognitive functions* (9 – 21). Oxford: Oxford University Press.

Baddeley, A., Della Salla, S., Gray, C., Papagano, C., & Spinnler, H. (1997). Testing central executive functioning with a pencil-and-paper test. In P. Rabbitt (ed.), *Methodology of frontal and executive function* (pp. 61 - 80). Hove, UK: Psychology Press.

Bartenwerfer, H. (1964). Allgemeine Leistungstests. In R. Heiss (Hrsg.). *Handbuch der Psychologie, Psychologische Diagnostik* (S. 385 – 410). Göttingen: Hogrefe.

- Bartenwerfer, H. (1983). *Allgemeine Leistungsdiagnostik*. In K. J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie, Serie II: Psychologische Diagnostik; Bd. 2: Intelligenz- und Leistungsdiagnostik* (S. 482 – 521).
- Bäumler, G. (1985). *Farb-Wort-Interferenztest FWIT nach J.R. Stroop*. Göttingen: Hogrefe.
- Beauducel, A., Wittmann, W.W. (2005). Simulation Study on Fit Indexes in CDA Based on Data With Slightly Distorted Simple Structure. *Structuring Equation Modeling* 2005; 41 – 75.
- Beckmann, J. (1991). Erhöhte Konzentration als Folge von Aufmerksamkeitsstörungen: Ein Zwei-Phasen-Modell. In J. Jansen, E. Hahn & H. Strang (Hrsg.), *Konzentration und Leistung*. Göttingen: Hogrefe.
- Beckmann, J. & Strang, H. (1993). Konzentration: Überlegungen zu einem vernachlässigten Konstrukt. In J. Beckmann, H. Strang & E. Hahn (Hrsg.), *Aufmerksamkeit und Engergetisierung* (S. 11 – 32). Göttingen: Hogrefe.
- Berg, A. & Westhoff, K. (2006). Facetten konzentrierten Arbeitens. *Report Psychologie* 31, Zeitschrift des Berufsverbands Deutscher Psychologinnen und Psychologen e.V.
- Berg, D. (1991). Zur Bedeutung von Reizkomplexität und –modalität bei Konzentrationsschwierigkeiten. *Zeitschrift für Pädagogische Psychologie*, 5, 9 – 20.
- Berg, D. & Erlwein, M. (1991). Gute Resultate im Aufmerksamkeits-Belastungs-Test nur aufgrund hoher Konzentrationsfähigkeit? *Psychologie in Erziehung und Unterricht* 38 (1991) S. 59 - 62.
- Berg, D. & Imhof, M. (1998). Aufmerksamkeit und Konzentration. In D. Rost (Hrsg.): *Handbuch der Pädagogischen Psychologie*. Weinheim: Beltz.

- Berg, D. & Imhof, M. (2001). Aufmerksamkeit und Konzentration. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. (2. Aufl.). Weinheim: Beltz.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Brähler, E., Holling, H., Leutner, D. & Petermann, F. (Hrsg.). (2002). *Brickenkamp Handbuch psychologischer und pädagogischer Psychologie*. (2. Aufl.). Weinheim: Beltz.
- Brickenkamp, R. (Hrsg.). (1975). *Handbuch psychologischer und pädagogischer Tests*. Göttingen: Hogrefe.
- Brickenkamp, R. & Karl, G. A. (1986). Geräte zur Messung von Aufmerksamkeit, Konzentration und Vigilanz. In R. Brickenkamp (Hrsg.), *Handbuch apparativer Verfahren in der Psychologie* (S. 195 – 211). Göttingen: Hogrefe.
- Brickenkamp, R. (1987). *Test d2. Aufmerksamkeits-Belastungs-Test*. 6. Aufl. Göttingen: Hogrefe.
- Brickenkamp, R. (2002). *Test d2: Aufmerksamkeits-Belastungs-Test* (9. überarb. und neu normierte Aufl.). Göttingen: Hogrefe.
- Bühner, M. (2001). Die Bedeutung von mentaler Operation und Material bei der Erfassung der Konzentrationstest-Leistung: Ein Beitrag zur Validität von Durchstreichetests. (1. Aufl.) Hamburg: Kovac.
- Bühner, M., Schmidt-Atzert, L., Grieshaber, E. & Lux, A. (2001). Faktorenstruktur verschiedener neuropsychologischer Tests – Ergebnisse einer retrospektiven Studie mit hirngeschädigten Patienten. *Zeitschrift für Neuropsychologie*, 12, 179 – 185.
- Bühner, M. & Schmidt-Atzert, L. (2002). Selbstbeurteilungen der Aufmerksamkeit: Ein Vergleich zwischen Hirngeschädigten und Gesunden. *Zeitschrift für Neuropsychologie*, 13, 263 - 269.

- Bühner, M. (2004). Einführung in die Test- und Fragebogenkonstruktion. Pearson-Studium.
- Bühner, M., Mangels, M., Krumm, S. & Ziegler (2005). Are Working Memory and Attention Related Constructs? *Journal of Individual Differences, Vol. 26 (3)*, 121 - 131.
- Bühner, M., Krumm, S., & Pick, M. (2005). Reasoning = working memory \neq attention. *Intelligence, 33*, 251 - 272.
- Bühner, M., Krumm, S., Ziegler, M. & Pluecken, T. (2006). Cognitive abilities and their interplay. Reasoning, crystallized intelligence, working memory capacity, and sustained attention. *Journal of Individual Differences, 27*, 57 - 72.
- Bühner, M. (2006). Einführung in die Test- und Fragebogenkonstruktion. 2. Aufl. Pearson-Studium
- Bühner, M., Ziegler, M. (2009). Statistik für Psychologen und Sozialwissenschaftler. Pearson-Studium.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523 - 547.
- Byrne, B., M. (2001). *Structural equation modeling with AMOS. Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Carlson, S. M. (2003). Executive Function in Context: Developmental, Measurement, Theory, and Experience. *Monographs of the Society for Research in Child Development 68*: 138 – 151.
- Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. New York: Cambridge University Press.

- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404 – 431.
- Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1 – 22.
- Cattell, R.B. & Horn, J. (1978). A check on the theory of fluid and crystallized intelligence with description of new sub-test designs. *Journal of Educational Measurement*, 15 (3), 139 – 164.
- Cattell, R.B. (1987). *Intelligence: Its structure, growth, and action*. Amsterdam: North-Holland.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. D., Botvinick, M. & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: who's in control? *Nature Neuroscience*, 3 (5): 421 - 3.
- Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, 19, 450 – 466.
- Debus, G. (1995). Konzentration. In W. Sarges (Hrsg.). *Management-Diagnostik*. (2. überarb. und erw. Aufl., S. 316 – 324). Göttingen: Hogrefe.
- Dries, C., Krumm, S., Lux, V. (2006). Prozessertifizierung der Führungskräfteauswahl nach DIN 33430 in einer deutschen Großbank. In: Westhoff, K. (Hrsg.) *Nutzen der DIN 33430. Praxisbeispiele und Checklisten* (S. 140), Lengerich: Pabst Science Publishers.

- Enders, P. (2004). Zur faktoriellen Struktur der Konzentrationsleistung. Ein Beitrag zur Konstruktvalidität von Konzentrationstests. Diplomarbeit, Philipps-Universität Marburg.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A., (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128, 309 - 331.
- Fisseni, H. J. (1997) Lehrbuch der psychologischen Diagnostik (2.Auf.). Göttingen: Hogrefe
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., De Fries, J. C., Hewitt, J. K. (2006) Not All executive functions are related to intelligence. *Psychological Science*, 2006, 17, No. 2, 172 – 179.
- Goldhammer, F. & Moosbrugger, H. (2006). Aufmerksamkeit. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 16 - 32). Heidelberg: Springer.
- Goldstein, E. B. (1989). *Sensation and perception* (3rd ed.). Pacific Grove: Brooks/Cole Publishing Company.
- Goschke, T. (2002). Volition und kognitive Kontrolle. In: Müsseler J, Prinz W. (Hrsg.) *Allgemeine Psychologie* (271 – 335). Heidelberg: Spektrum.
- Grant, D. A. & Berg, E. A. (1981). WCST, Wisconsin Card Sorting Test.
- Griffith, R. L., Chmielowski, T., Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behaviour. *Personnel Review*, Vol. 36 (3), pp. 341 - 355.

- Gustafsson, J. E. (1988). Hierarchical models of individual differences in cognitive abilities: In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence, Vol. 4*, Hillsdale, New Jersey: Lawrence Erlbaum.
- Häcker, H. & Stapf, K. H. (1998). *Dorsch Psychologisches Wörterbuch*. (13. überarb. u. erw. Aufl.). Bern: Huber.
- Heyde, G. (2000). INKA Inventar Komplexer Aufmerksamkeit. (2. Aufl.). Frankfurt: Swets Test Services.
- Hoffmann, J. (1993). Konzentration durch Antizipation. In Beckmann, J., Strang, H. & Hahn, E. (Hrsg.): *Aufmerksamkeit und Energetisierung* (S. 35-66). Göttingen: Hogrefe.
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6, 65 - 70.
- Horn, J.L. (1988). Thinking about human abilities. in: J.R. Nesselroade & R. B. Cattell (Eds.). *Handbook of multivariate experimental psychology*, 645 – 685. New York: Plenum Press.
- Hu, L.T., Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* 1999; 6: 1 – 55.
- Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). Berliner Intelligenzstruktur-Test BIS-Test, Form 4. Göttingen: Hogrefe
- Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, Whole No. 89.

- Jonides, J., & Smith, E. E. (1997). The architecture of working memory. In M. D. Rugg (Ed.), *Cognitive neuroscience* (pp. 243 – 276). Cambridge, MA: MIT Press.
- Kersting, M., Hornke, L. (2003). Qualitätssicherung und –optimierung in der Diagnostik: die DIN 33430 und notwendige Begleit- und Folgeinitiativen. *Psychologische Rundschau*, 54 (3), 175 – 184. Göttingen: Hogrefe.
- Kersting, M. (2006). Zur Beurteilung der Qualität von Tests: Resümee und Neubeginn. *Psychologische Rundschau*, 57 (4), 243 – 253.
- Kristof, W. (1963). *Die Verteilung aufgewerteter Zuverlässigkeitskoeffizienten auf der Grundlage von Testhälften*. *Archiv für die gesamte Psychologie*, 115, 230 - 240.
- Kröner, S. (2001). *Intelligenzdiagnostik per Computersimulation*. Münster: Wasmann.
- Lehrl, S. (1995). MWT-B Mehrfachwahl-Wortschatz-Intelligenz-Test. (3. überarb. Aufl.) Balingen: Perimed-Spitta Verlag.
- Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology*, 49A, 29 – 50.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. (6. Aufl.). Weinheim: Beltz.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83 (404), 1198 – 1202.
- Little, R. J., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (2nd ed.). Hoboken, NJ: Wiley.

- Little, T.D., Cunningham, W.A., Shahar, G. & Widaman, K.F. (2002). To Parcel or Not to Parcel: Exploring the Question, Weighing the Merits. *Structural Equation Modeling*, 9 (2), 151 – 173.
- Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189 – 239). San Diego, CA: Academic Press.
- Logan, G. D. & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, 108, 393 - 434.
- Logan, G. D. & Bundesen, C. (2003). Clever Homunculus: Is There an Endogenous Act of Control in the Explicit Task-Cuing Procedure? *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 29, No. 3, 575 - 599.
- Logan, G. D. (2004). Working Memory, Task Switching, and Executive Control in the Task Span Procedure. *Journal of Experimental Psychology: General* 2004, Vol 133, No. 2, 218 – 236.
- Luciano, M., Wright, M. J., Smith, G. A., Geffen, G. M., Geffen L. B., & Martin, N. G. (2001). Genetic covariance among measures of information processing speed, working memory, and IQ. *Behaviour Genetics*, 31, 581 – 592.
- Lüdtke, O., Robitzsch, A., Trautwein, U., Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, Vol. 58 (2), 103 - 117.
- Lukesch, H. & Mayrhofer, S. (2001). *KLT Konzentrations-Leistungs-Test revidierte Fassung* (1. Aufl.). Göttingen: Hogrefe.

- Marschner G (1972/1980). Revisions-Test nach Dr. B. Stender (2. Aufl.). Göttingen: Hogrefe
- Mierke, K. (1957). *Konzentrationschwäche*. Bern: Huber.
- Miller, E. K. & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 24: 167 - 202.
- Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442 - 481). New York: Cambridge Univ. Press.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41, 49 – 100.
- Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 93 – 148). Hove, UK: Erlbaum.
- Monsell, S., & Driver, J. (2000). Banishing the control homunculus. In: Monsell S, Driver J. (Hrsg). *Control of cognitive processes* (3 – 32). Cambridge: The Mit Press.
- Moosbrugger, H. & Oehlschlägel, J. (1996). *Frankfurter Aufmerksamkeits-Inventar*. Göttingen: Hogrefe.
- Moosbrugger, H. & Heyden, M. (1998). *Frankfurter Adaptiver Konzentrationsleistungs-Test (FAKT)*. Programmversion Hogrefe TestSystem 3.1.0 1999. Bern: Huber
- Moosbrugger, H., Goldhammer, F., Schweizer, K. (2006). Latent Factors Underlying Individual Differences in Attention Measures. *European Journal of Psychological Assessment*, Vol 22 (3): 177 – 188.

- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*, 111 - 121.
- Nettelstroth, W. (2003). *Intelligenz im Rahmen der beruflichen Tätigkeit. Unveröffentlichte Dissertation, Freie Universität Berlin.*
- Neubauer, A. C. & Bucik, V. (1996). The mental speed-IQ relationship: unitary or modular? *Intelligence*, *22*, 23 – 48.
- Neubauer, A. (1993). Intelligenz und Geschwindigkeit der Informationsverarbeitung: Stand der Forschung und Perspektiven. *Psychologische Rundschau*, *44*, 90 - 105.
- Neumann, O. & Simon, U. (1994). Bewusstsein, Aufmerksamkeit, Konzentration. *Psychologie und Sport*, *4*, 123 – 139.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. Consciousness and self-regulation. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1 – 18). New York: Plenum.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266 - 300.
- Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W.W. (1999). *The multiple facets of working memory – Storage, processing, supervision, and coordination.* Mannheim: Universität Mannheim, Lehrstuhl für Psychologie II.

- Oberauer K., Süß H.-M., Schulze R., Wilhelm O., Wittmann W.W. (2000). Working memory capacity – facets of a cognitive ability construct. *Personality and Individual Differences*, 29, 1017 – 1045.
- Oberauer, K., Süß, H.-M., Wilhelm, O., Wittmann, W.W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, 31, 167 – 193.
- Oberauer, K., Süß, H.-M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. N. Towse (Eds.), *Variation in working memory* (pp. 49 - 75). New York: Oxford University Press.
- Oehlschlägel, J., & Moosbrugger, H. (1991). Überraschende Validitätsprobleme im Aufmerksamkeits-Belastungs-Test d2. *Report Psychologie*, September 1991, 16 - 25.
- Oswald, W. D. & Roth, E. (1987). ZVT Der Zahlen-Verbindungs-Test. Göttingen: Hogrefe.
- Phillips, L. H. (1997). Do "frontal tests" measure executive function? Issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 191 - 213). Hove, UK: Psychology Press.
- Posner, M. I. (1980). Orientation of attention. *Quarterly Journal of Experimental Psychology*, 32, 3 - 25.
- Posner, M. I. & Rafal, R. D. (1987). Cognitive theories of attention and the rehabilitation of attention deficits. In M. J. Meier, A. L. Benton & L. Diller (Eds.), *Neuropsychological Rehabilitation* (pp. 182 – 201). Edinburgh: Churchill Livingstone.

Posner, M. L., & Raichle, M. E. (1994). *Images of mind*. New York: Sci. Am.

Rabbitt, P. (Ed.). (1997a). *Methodologies of frontal and executive function*. Hove, UK: Psychology Press.

Rabbitt, P. (1997b). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1 - 38). Hove, UK: Psychology Press.

Rapp, G. (1982). *Aufmerksamkeit und Konzentration. Erklärungsmodelle – Störungen – Handlungsmöglichkeiten*. Bad Heilbrunn: Klinhardt.

Raven, J.C. (1976). *Advanced progressive matrices, Sets I and II*. Oxford, England: Oxford Psychologists Press.

Rindermann, H., Neubauer, A.C. (2004). Processing speed, intelligence, creativity and school performance: Testing of causal hypotheses using structural equation models. *Intelligence*, 32, 573 - 589.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, 123, 374 - 393.

Roberts, A. C., (2003). Introduction. In: Roberts A.C., Robbins T.W., Weiskrantz, L. (Hrsg). *The prefrontal cortex. Executive and cognitive functions* (1 – 5). Oxford: Oxford University Press.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207 – 231.

- Samstag, K., Sander, A., Schmidt, R. & Ingenkamp, K. (Hrsg.). (1971). *Diagnostischer Rechentest für 3. Klassen DRE 3*. Weinheim/Berlin: Beltz.
- Schneider, W. X., Owen, A.M., Duncan, J. (2000). Executive control and the frontal lobe: Current issues: Berlin: Springer.
- Schmidt-Atzert, L. & Bühner, M. (1997). Was misst der d2-Test? Eine experimentelle Analyse von Testleistungen. *Diagnostica*, 43, 314-326, Hogrefe: Göttingen.
- Schmidt-Atzert, L., & Bühner, M. (1998). Fehlertypen im Aufmerksamkeits-Belastungs-Test d2. *Diagnostica*, 44, 142 – 152.
- Schmidt-Atzert, L. (2002). Tests und Tools. Intelligenz-Struktur-Test 2000 R. *Zeitschrift für Personalpsychologie*, 1, 50 – 56.
- Schmidt-Atzert, L., Büttner, G., Bühner, M. (2004). Theoretische Aspekte von Aufmerksamkeits-/Konzentrationsdiagnostik. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit*. Jahrbuch der pädagogisch-psychologischen Diagnostik, S. 3 – 22 Göttingen, Hogrefe.
- Schmidt-Atzert, L., Bühner, M., Rischen, S., Warkentin, V. (2004). Erkennen von Simulation und Dissimulation im Test d2. *Diagnostica*, 50, Heft 3, 124 - 133. Hogrefe: Göttingen.
- Schmidt-Atzert, L., Bühner, M. & Enders, P. (2006). Messen Konzentrationstests Konzentration? Eine Analyse der Komponenten von Konzentrationsleistungen. *Diagnostica*, 52, Heft 1, 33 - 44.
- Schmidt-Atzert, L., Krumm S., Bühner M. (2008). Aufmerksamkeitsdiagnostik. Ableitung eines Strukturmodells und systematische Einordnung von Tests. *Zeitschrift für Neuropsychologie*, 19 (2) 59 - 82.

- Schweizer, K. (1995a). Hypothesen zu den biologischen und kognitiven Grundlagen der allgemeinen Intelligenz. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 16, 67 – 81.
- Schweizer, K. (1995b). Kognitive Korrelate der Intelligenz. In D. Albert, L. Pawlik, K. H. Stapf & W. Stroebe (Hrsg.), *Lehr- und Forschungstexte der Psychologie*. Göttingen: Hogrefe.
- Schweizer, K. & Moosbrugger, H. (1999). Aufmerksamkeit, Intelligenz und Verarbeitungsgeschwindigkeit als Komponenten der mentalen Leistungsfähigkeit. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 20, 126 - 132.
- Shiffrin, R.-M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review* 85 (2): 127 – 190.
- Simon, H. A. (1975) TOH, Tower of Hanoi.
- Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology*, 89, 669 – 679.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643 – 662.
- Sturm, W. (1997). Neuropsychologische Diagnostik. In W. Hartje & K. Poek (Hrsg.), *Klinische Neuropsychologie* (S. 59 - 79). Stuttgart: Thieme.
- Sturm, W. (2002). Kognitive Kontrolle der Aufmerksamkeitsintensität: Funktionelle Neuroanatomie. *Zeitschrift für Psychologie*, 212, (2), 107 – 114.
- Sturm, W. (2004). Kognitive Kontrolle der Aufmerksamkeitsintensität. *Zeitschrift für Psychologie*, 212 (2), 107 – 114. Göttingen: Hogrefe.

- Süß, H.-M. (1995). *Intelligenz, Wissen und Problemlösen*. Habilitationsschrift, Fakultät für Philosophie, Psychologie und Erziehungswissenschaft der Universität Mannheim.
- Süß, H.-M., Oberauer, K., Wittmann, W.W., Wilhelm, O. & Schulze, R. (1996). *Working memory capacity and intelligence: an integrative approach based on Brunswik Symmetry* (Report 8). Mannheim: Universität Mannheim, Lehrstuhl Psychologie II.
- Tewes, U. (1991). *Hamburg-Wechsler-Intelligenztest für Erwachsene Revision 1991*. Bern: Huber.
- Van Zomeren, A.H. & Brouwer, W.H. (1994). *Clinical Neuropsychology of Attention*. New York: Oxford Press.
- Vickers, D., Nettelbeck, T., & Wilson, R. J. (1972). Perceptual indices of performance, the measurement of inspection time and noise in a visual system. *Perception, 1*, 263 – 295.
- Vickers, D., & McDowell, A. (1996). Accuracy in the frequency accural speed test (FAST), Inspection time and psychometric intelligence in a sample of primary school children. *Personality and Individual Differences, Vol. 20, No. 4*, 463 - 469.
- Waltz, J.A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10*, 119 - 125.

- West, S.G., Finch, J.F. & Curran, P.J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R.H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (S. 56 – 75). Thousand Oaks, CA: Sage Publications.
- Westhoff, K. & Kluck, M. L. (1983). Zusammenhang zwischen Intelligenz und Konzentration. *Diagnostica*, 29, 310 – 319.
- Westhoff, K., & Kluck, M. L. (1984). Ansätze einer Theorie konzentrativer Leistungen. *Diagnostica*, 3, 167 – 183.
- Westhoff, K. (1989). Übungsabhängigkeit von Leistungen in Konzentrationstests. *Diagnostica*, 35, 122 – 130.
- Westhoff, K. & Dewald, D. (1990). Effekte der Übung in der Bearbeitung von Konzentrationstests. *Diagnostica*, 36, 1 – 15.
- Westhoff, K. (1991). Das Akku-Modell der Konzentration. In H. Barchmann, W. Konze und N. Roth (Hrsg.): *Aufmerksamkeit und Konzentration im Kindesalter* (S. 47-55). Berlin: Verlag Gesundheit.
- Westhoff, K. & Hagemester, C. (1992). Reliabilität von Fehlern in Konzentrationstests. *Diagnostica*, 38, 116 – 129.
- Westhoff, K. (1995). Aufmerksamkeit und Konzentration. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie des Bereiches Differentielle Psychologie, Band 2 der Bereiche/Dimensionen individueller Differenzen* (S. 375 – 402), Göttingen: Hogrefe.
- Wittmann, W. W. (1987). Grundlagen erfolgreicher Forschung in der Psychologie: Multimodale Diagnostik, Multiplismus, multivariate Reliabilitäts- und Validitätstheorie. *Diagnostica*, 33, 3, 209 – 226.

- Wittmann, W. W. (1988). Multivariate reliability theory: Principles of symmetry and successful validation strategies. In J. R. Nesselroade & R. B. Cattell (Hrsg.); *Handbook of multivariate experimental psychology* (2. Aufl.). New York: Plenum.
- Wylie, G., & Allport, A. (2000). Task switching and the measurement of „switch costs“. *Psychological Research*, 63, 212 – 233.
- Ziegler, M., Schmidt-Atzert, L., Bühner, M., Krumm, S. (2007). Fakability of different measurement methods for achievement motivation: Questionnaire, semi-projective, and objective. *Psychology Science, Vol. 49 (4)*, pp. 291 - 307.
- Zimmermann P. & Fimm B. (1993). Testbatterie zur Aufmerksamkeitsprüfung (TAP), Version 1.0. Handbuch. Würselen: Vera Fimm / Psychologische Testsysteme sowie (1994) Version 1.02c und (2002) Version 1.7.
- Zimmermann, P. & Fimm, B. (2002). A test battery for attentional performance. In: M. Leclercq & P. Zimmermann (eds). *Applied Neuropsychology of Attention. Theory, Diagnosis and Rehabilitation*. pp. 110 - 151.

Validierungsstudie 1

Als Validierungsstudie 1 diente ein Datensatz, der aus einer Erhebung zweier Studentinnen der Philipps-Universität Marburg stammt.

1 Testdurchführung

Die Erhebung der Daten fand nach einem weitgehend festgelegten Schema im September und Oktober 2002 statt. Grundsätzlich fand die Untersuchung als Gruppentestung mit je fünf Teilnehmern statt. Einige Probanden konnten aufgrund technischer Schwierigkeiten (z.B. unausgeglichene Hörgeräte oder Computerabstürze) nicht alle Tests vollständig durchführen, was letztlich zu teilweise unvollständigen Testdaten führte.

2 Stichprobe

Der Datensatz umfasste ursprünglich 139 Personen, von denen 1 Fall aufgrund extremer Ausreißerwerte in den TEMEKKO Variablen eliminiert wurde. Der verbliebene Datensatz umfasste 138 Personen im Alter von 59 bis 92 Jahren ($M = 69.5$ Jahre; $SD = 7.17$ Jahre). Das mittlere Alter der 89 weiblichen Versuchspersonen (64.5 %) betrug 70.1 Jahre ($SD = 7.59$ Jahre), das der 49 männlichen Probanden 68.6. Jahre ($SD = 6.29$ Jahre).

Zum Zeitpunkt der Datenerhebung waren die Hälfte der Personen ($N = 69$) verheiratet, 36 waren verwitwet, 20 ledig und 7 geschieden. Eine Person war zum Erhebungszeitpunkt getrennt lebend. 35.5 % der Versuchspersonen ($N = 49$) waren Akademiker, 41 Personen

verfügten über einen Hauptschulabschluss, 34 Versuchspersonen hatten Realschulabschluss und 11 Personen waren Abiturienten. 3 Personen machten keine Angaben. Trotz des fortgeschrittenen Alters der Probanden waren 21 Personen (noch) berufstätig. 5 Personen machten hierzu keine Angaben. Die berufliche Tätigkeit der Personen umfasste die Funktionen nichtleitender Angestellter ($N = 28$), Beamter des höheren oder gehobenen Dienstes ($N = 26$), leitender Angestellter ($N = 22$) oder Hausfrau/-mann ($N = 21$). 37 Personen waren Selbständiger (mit oder ohne Mitarbeiter), Facharbeiter, Beamte des einfachen oder mittleren Dienstes oder Arbeitnehmer ohne Berufsausbildung. 4 Personen machten hierzu keine Angaben; wobei sich die genannten Angaben nicht nur auf den momentanen Status bezogen, sondern auch die Zeit vor der Pensionierung einschloss. Nahezu die Hälfte aller Versuchspersonen dieser Stichprobe ($N = 62$) gaben an, noch nie am Computer gearbeitet zu haben. Jeweils 23 Personen taten dies ein- oder mehrmals in den letzten Jahren oder pro Woche. Mehr als 11 %, nämlich 16 Personen, waren ein- oder mehrmals täglich am Computer tätig, nur 13 Personen ein- oder mehrmals im Monat. 2 Personen machten keine Angaben.

Die Form der Anwerbung der Versuchspersonen sowie Informationen über Rückmeldungen oder Vergütung der Testung lagen nicht vor. Als Selektionskriterium galt, dass die Teilnehmer bereits im fortgeschrittenen Alter sein und an keiner aufmerksamkeits einschränkenden neurologischen Erkrankung leiden sollten.

3 Versuchsmaterial

Nachfolgende Tabelle zeigt den chronologischen Ablauf der Testung. Die Gesamtdauer der Untersuchung belief sich auf ca. 3.5 Stunden. Die Testverfahren Aufmerksamkeitsfragebogen, OLMT, der Test d2 groß sowie der Subtest

Gesichtsfeldprüfung aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP) werden der Vollständigkeit halber gelistet. Die Daten aus diesen Verfahren flossen aber nicht in die vorliegende Validierungsuntersuchung zum TEMEKKO ein.

Tabelle I.1: Ablaufplan der Testung

Testverfahren	Untertest	Dauer	Art der Bearbeitung
Aufmerksamkeitsfragebogen		n.b. ¹	Paper-Pencil
Wiener Testsystem, objektiver Leistungsmotivationstest (OLMT)		n.b. ¹	PC-gestützt
d2 Aufmerksamkeits- Belastungs- Test	klein	n.b. ¹	Paper-Pencil
TEMEKKO		14 min.	Paper-Pencil
<i>Pause</i>		15 min.	
d2 Aufmerksamkeits- Belastungs- Test	groß		Paper-Pencil
Testbatterie zur Aufmerksamkeitsprüfung (TAP)	Gesichtsfeldprüfung	n.b. ¹	PC-gestützt
	Alertness	n.b. ¹	PC-gestützt
	Visuelles Scanning (100 Trials)	n.b. ¹	PC-gestützt
	Geteilte Aufmerksamkeit (Serie 4, Test C, 60 Trials)	n.b. ¹	PC-gestützt
	Go/Nogo (Bedingung 2)	n.b. ¹	PC-gestützt

Tabelle I.1 wird fortgesetzt

Testbatterie zur Aufmerksamkeitsprüfung (TAP)	Arbeitsgedächtnis (Schwierigkeitsstufe 3, 100 Trials)	n.b. ¹	PC-gestützt
	Reaktionswechsel, verbale Version, komplexe Variante	n.b. ¹	PC-gestützt

¹ nicht berichtet

3.1 TEMEKKO

Auf die nähere Testbeschreibung des TEMEKKO wird an dieser Stelle verzichtet und auf Abschnitt 3 (Entwicklung des TEMEKKO) im Manual verwiesen.

3.2 Aufmerksamkeits-Belastungs-Test d2

(Brickenkamp, 1987). Der Test d2 ist ein standardisierter Konzentrationstest, der auf dem Durchstreicheprinzip nach Bourdon beruht. Er existiert in zweierlei Formaten (klein und groß). Der üblicherweise eingesetzte Test d2 ist jener im Format A4 (klein). Bei dem großen Test d2 handelt es sich lediglich um eine nach A3 Format vergrößerte Fassung. Die für die Validierung des TEMEKKO verwendete Fassung ist der Test d2 klein. Der d2 Aufmerksamkeits-Belastungstest gilt als Maß für selektive Aufmerksamkeit (oder sustained attention). Er misst die Schnelligkeit und Genauigkeit der Unterscheidung ähnlicher visueller Reize (Brickenkamp 1975 bzw. 1987). Der Test d2 besteht aus 14 Zeilen mit je 47 Buchstaben. Die Buchstaben (*d* oder *p*) haben null, einen oder zwei vertikale Striche über oder unter dem Zeichen, oder an beiden Stellen. Die Probanden müssen jene *d*'s mit insgesamt zwei Strichen innerhalb von 4 Minuten und 40 Sekunden durchstreichen. Alle anderen Kombinationen von Buchstaben und Strichen gelten als Non-Targets. Als Leistungsmaß galt der KL-Wert (Konzentrationsleistung), der sich aus der

Anzahl der richtigerweise durchgestrichenen Zeichen abzüglich der Verwechslungsfehler berechnet.

3.3 Testbatterie zur Aufmerksamkeitsprüfung (TAP)

(Zimmermann & Fimm, 1993). Die TAP ist ein vollständig computergestütztes Verfahren. Für einzelne Untertests werden separate Tasten an den Computer angeschlossen, über die die Versuchsperson die jeweiligen Aufgaben beantwortet. Für Tests mit akustischen Darbietungen werden Kopfhörer an die Soundkarte des Computers angeschlossen. Zu Beginn jedes Untertests wird die Instruktion auf dem Bildschirm dargestellt. Daraufhin hat der Proband die Möglichkeit, sich in einem Vorversuch mit den spezifischen Aufgabenanforderungen vertraut zu machen. So kann sichergestellt werden, dass die Instruktion richtig verstanden wurde. Bei Verständnisproblemen kann der Vorversuch beliebig oft wiederholt werden. Erst danach erfolgt die auszuwertende Hauptaufgabe. Die Auswertung der einzelnen Subtests erfolgte durch das Testprogramm und konnte direkt in eine SPSS-Datei umgewandelt werden.

Aus der Testbatterie zur Aufmerksamkeitsprüfung wurden die Untertests Alertness, Geteilte Aufmerksamkeit, Go/Nogo, Visuelles Scanning, Arbeitsgedächtnis und Reaktionswechsel verwendet. Diese Untertests sollen die beiden Kernkomponenten der Aufmerksamkeit im Modell bei van Zomeren und Brouwer (1994), den Selektivitäts- bzw. Intensitätsaspekt, erfassen. Die TAP basiert auf dem Ansatz, dass Aufmerksamkeit keine einheitliche Funktion darstellt, sondern sich aus mehreren spezifischen Teilfunktionen zusammensetzt, die die Kontrolle des Informationsflusses im kognitiven System beschreiben. Die Verfahren der TAP operationalisieren einfache Reaktionsparadigmen, bei

denen selektiv auf gut diskriminierbare, sprachfreie Reize mittels Tastendruck zu reagieren ist.

3.3.1 Alertness

Dieser Untertest erfasst die Reaktionszeit auf einen kritischen (visuellen) Reiz mit einem oder ohne einen vorangehenden akustischen Warnreiz (Ton). Somit kann zum einen die allgemeine Reaktionsschnelligkeit (tonische Alertness) erfasst werden und zum anderen die Fähigkeit, in Erwartung eines Reizes das Aufmerksamkeitsniveau zu steigern und aufrecht zu erhalten (phasische Alertness).

Bei der Aufgabe erscheint in der Mitte des Bildschirms ein Kreuz, auf das, in der Bedingung ohne Warnreiz, so schnell als möglich bei Erscheinen des Kreuzes durch Tastendruck reagiert werden soll. Bei der Bedingung mit Warnreiz ertönt vor dem Erscheinen des Kreuzes ein akustischer Reiz (Ton). Reagiert werden soll wiederum auf das Aufscheinen des visuellen Reizes (Kreuz). Der Test besteht aus vier Testblöcken (zu je 20 Reizen), die in einem ABBA-Design konzipiert sind. (A = ohne Warnreiz, B = mit Warnreiz). Bleibt ein Reiz länger als zwei Sekunden unbeantwortet, oder erfolgt ein Tastendruck bereits bei Ertönen des akustischen Reizes ($RZ < 100$ ms), werden diese Reize als unbeantwortet erfasst und wiederholt. Das Intervall zwischen den einzelnen Reizen variiert zwischen 300 ms und 700 ms. Als Leistungsmaße dienen die Mediane der Reaktionszeiten ohne Warnton (tonische Alertness).

3.3.2 Geteilte Aufmerksamkeit

Es handelt sich um eine Dual-Task Aufgabe, d.h. es müssen gleichzeitig zweierlei Reizdarbietungen beachtet werden. Die Teilung der Modalitäten (optisch und akustisch) soll sicherstellen, dass Interferenzen zwischen den Informationskanälen unterbunden werden. Bei der optischen Aufgabe erscheint eine 4 x 4 Punkte-Matrix auf dem Bildschirm. Einige dieser Punkte werden in zufälliger Anordnung durch Kreuze ersetzt. Sobald vier Kreuze ein Quadrat bilden, ist so schnell als möglich durch Tastendruck zu reagieren. Die akustische Aufgabe (über Kopfhörer) umfasst die Darbietung von alternierenden Tönen, jeweils einem hohen und einem tiefen Ton. Sobald zwei identische Töne nach einander wahrgenommen werden, ist so schnell als möglich die Taste zu drücken. Insgesamt werden 100 optische und 200 akustische Aufgaben dargeboten; aus denen jeweils 15 optische und 15 akustische Zielreize durch Tastendruck zu identifizieren sind. Als Leistungsmaß dient der Median der Reaktionszeiten. Für die vorliegende Untersuchung wurde aus der Serie 4 der Test 3 (Quadrate und Töne) dargeboten; die Mediane der Reaktionszeiten der optischen Aufgabe (Quadrate) wurde für die Datenanalyse verwendet.

3.3.3 Go/Nogo

Der Test erfasst die Fähigkeit zur Reaktionsunterdrückung bei irrelevanten Reizen und die Reaktionszeit unter Reizselektionsbedingungen (selektive Aufmerksamkeit). Den Versuchspersonen werden fünf Quadrate (3 x 3 cm) mit unterschiedlichen Füllmustern präsentiert, sowie zwei dieser fünf Reize, die als kritische Reize definiert sind (Testbedingung 2). Nach einer kurzen Einprägephase werden in der Mitte des Bildschirms

nacheinander die einzelnen Quadrate mit Füllmustern in zufälliger Abfolge präsentiert. Bei Erscheinen eines kritischen Reizes muss so schnell als möglich per Tastendruck reagiert werden. Insgesamt werden 60 Items dargeboten, die 24 Zielreize enthalten. Als Leistungsmaß diente der Median der Reaktionszeiten.

3.3.4 Visuelles Scanning

Bei diesem Untertest wird dem Probanden auf dem Bildschirm eine 5 x 5 Matrix mit Quadraten dargeboten, die jeweils an einer Seite offen sind. Am linken Rand des Bildschirms erscheint der kritische Reiz. Es ist ein Quadrat, das nach einer Seite offen ist und nach dem in der vorgegebenen Matrix gesucht werden soll. Durch Tastendruck antwortet der Proband, ob das Target in der Reizvorlage enthalten oder nicht enthalten ist. Insgesamt werden 100 Reizvorlagen dargeboten, von denen 50 % Targets enthalten. Als Leistungsmaß diente der Median der Reaktionszeiten der kritischen Trials.

3.3.5 Arbeitsgedächtnis

In der Schwierigkeitsstufe 3 dieses Untertests werden dem Probanden nacheinander einstellige Zahlen präsentiert. Sobald eine Zahl gleich ist mit der vorletzten Präsentierten, ist eine Reaktionstaste zu drücken. Aus 100 dargebotenen Reizen ist auf 15 Targets zu reagieren. Als Leistungsmaße wurden die Mediane der Reaktionszeiten erfasst.

3.3.6 Reaktionswechsel

In diesem Test werden dem Probanden gleichzeitig links und rechts vom Fixationspunkt zwei konkurrierende Reize (ein Buchstabe und eine Zahl) dargeboten. Die Aufgabe besteht darin, die Taste auf der Seite zu drücken, auf der sich das Target befindet. Das Target ist alternierend einmal der Buchstabe, einmal die Zahl. Insgesamt werden 100 Reize dargeboten. Die Mediane der Reaktionszeiten galten als Leistungsmaße.

4 Statistische Analysen

4.1 Deskriptiva

Nachfolgende Tabelle I.2 zeigt einen Überblick über die Rohdaten der Analysevariablen.

Tabelle I.2: Deskriptive Statistiken der Testvariablen

Tests	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>	<i>r_{age}</i>	<i>r_{gender}</i>
Te fig	138	6.0	53.0	30.7	8.3	-0.1	0.1	-.41**	.24**
Te num	137	7.0	59.0	31.6	9.8	0.4	0.2	-.33**	.20*
Te verb	137	9.0	54.0	32.5	8.3	-0.0	0.2	-.38**	.05
Te g	137	35.0	150.0	94.8	22.5	-0.0	0.2	-.44**	.19*
d2	134	51.0	270.0	149.5	38.2	0.1	0.2	-.27**	.01
ga	135	643.0	1,299.0	931.4	127.6	0.7	0.4	.20*	-.02
rw	130	642.0	4,372.5	1,297.2	542.1	2.2	0.2	.41**	.08
ag	132	324.5	1,603.0	701.9	208.2	1.3	0.2	.10	-.12
vs	138	1,695.5	8,601.0	2,922.6	1,030.9	2.2	0.2	.49**	-.07
gn	138	455.0	858.5	612.2	74.8	0.8	0.2	.17	-.06
al	135	179.0	558.0	286.8	62.8	1.3	0.2	.34**	-.27**

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO; d2 = Test d2; ga = geteilte Aufmerksamkeit, Quadrate, TAP; rw = Reaktionswechsel, TAP; ag = Arbeitsgedächtnis, TAP; vs = Visuelles Scanning, TAP; gn = Go/Nogo, TAP; al = Alertness, TAP. *r* = bivariate Korrelation, * = $p < .05$, ** = $p < .01$.

Die augenscheinliche Inspektion der Verteilungsformen der einzelnen Testvariablen anhand der Histogramme und zugehörigen Normalverteilungskurven erbrachte keine Auffälligkeiten. Lediglich die Variablen Reaktionswechsel und Visuelles Scanning zeigten stark linkssteile Verteilungen (auch hinsichtlich der numerischen Kennwerte der Schiefe), was aufgrund der Variablen als Reaktionszeiten erklärbar ist. Die kreuzweise angefertigten Streudiagramme gaben nach optischer Prüfung Anlass, einen Fall aus dem Datensatz zu eliminieren. Auch wenn einige weitere Fälle auffällige Werte zeigten, so lagen sie dennoch im Trend. Es wurde daher lediglich ein Fall für die weiteren Berechnungen ausgeschlossen.

4.2 Zusammenhang mit Alter und Geschlecht

Die Korrelationen der z-transformierten und zum Teil (Reaktionszeiten) umgepolten Testvariablen mit den Variablen Alter und Geschlecht zeigten insbesondere mit Alter substantielle Zusammenhänge. Um den Einfluss von Alter und Geschlecht aus den Analysevariablen zu eliminieren, wurden die z-transformierten und ggfs. umgepolten Testkennwerte als abhängige Variable und Alter und Geschlecht als unabhängige Variablen einer linearen multiplen Regressionsanalyse (Einschluss Methode) unterzogen. Nachfolgende Tabelle zeigt die korrelativen Zusammenhänge der Analysevariablen vor und nach Auspartialisierung der Variablen Alter und Geschlecht. Alle Korrelationen wurden auf dem 1 % Niveau signifikant. Die optische Prüfung der Zusammenhänge mittels Streudiagrammen zeigte erwartungsgemäß durchgängig einen positiven linearen Trend.

Tabelle I.3: Zusammenhang z-transformierte Testvariablen und standardisierte Residuen

z \ Res.	Te fig	Te num	Te verb	Te g	d2	ga	rw	ag	vs	gn	al
Te fig	.899										
Te num		.935									
Te verb			.933								
Te g				.897							
d2					.968						
ga						.981					
rw							.906				
ag								.990			
vs									.871		
gn										.991	
al											.922

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te ges = TEMEKKO; d2 = Test d2; ga = geteilte Aufmerksamkeit, TAP; rw = Reaktionswechsel, TAP; ag = Arbeitsgedächtnis, TAP; vs = Visuelles Scanning, TAP; gn = Go/Nogo, TAP; al = Alertness, TAP.

4.3 Schätzung fehlender Werte

Wie im Manual unter Kapitel 4.1.2 (Prozedere der Datenbearbeitung) beschrieben wurde im Falle fehlender Werte der EM-Algorithmus angewendet. Der MCAR-Test (missing completely at random) nach Little war mit $\chi^2[308] = 453.23$, $p = < .001$ signifikant. Dessen ungeachtet wurden, um eine Verringerung der Stichprobengröße durch listenweisen Fallausschluss zu vermeiden, fehlende Werte durch Schätzung mittels EM-Algorithmus ergänzt. Die grafische Überprüfung (anhand von Streudiagrammen) der standardisierten Residuen mit fehlenden Werten und jenen Variablen, die um die fehlenden Werte ergänzt wurden sowie der korrelativen Zusammenhänge fiel zufriedenstellend aus. Bei $N = 138$ Fällen beinhalteten zwar alle Analysevariablen fehlende Werte, diese umfassten aber lediglich 1 (Minimum) bis 8 (Maximum) Fälle, i.e. max. 5.8 %.

4.4 Prüfung auf Gleichheit der Testformen und Materialarten

Das zu Grunde gelegte Messmodell (vgl. Abbildung 3 in Abschnitt 4.1.2 (Prozedere der Datenbearbeitung) im Manual) beinhaltet die Modellierung beider Fragestellungen und wurde zum einen für die Tempowerte (Modell a) und zum anderen für die Fehler-Prozent-Werte (Modell b) des TEMEKKO verwendet. Zunächst sei berichtet, dass die Annahme einer multivariaten Normalverteilung der Daten abgelehnt wurde (Tempowerte: multivariate Kurtosis = 5.77, c.r. = 3.46, $p = < .001$ bzw. Fehler-Prozent-Werte: multivariate Kurtosis = 38.57, c.r. = 23.12, $p = < .001$). Deshalb wurde eine Bollen-Stine-Bootstrap-Korrektur verwendet. Die Angaben der einzelnen Testvariablen hinsichtlich der von West, Finch und Curran (1995) postulierten Grenzen für Schiefe (< 2) und Kurtosis (< 7) waren für die Tempowerte durchgängig akzeptabel. Bei den Fehler-Prozent-Variablen waren die Grenzen der Variablen numerisch-Fehler-B-Prozente mit 3.47 für die Schiefe und 17.57 für die Kurtosis überschritten. Die Werte der Variable figural-Fehler-B-Prozente waren hinsichtlich der Schiefe mit 2.21 knapp über- und mit 6.96 knapp unterschritten.

Die Strukturgleichungsmodelle wiesen sowohl bei der Prüfung auf Passung der Tempowerte (Modell a) ($\chi^2 [16] = 27.29$, $p = \text{n.s.}$, $CFI = .98$, $RMSEA = .07$ [.02; .12], $SRMR = .07$) als auch der Fehler-Prozent-Werte (Modell b) ($\chi^2 [16] = 20.71$, $p = \text{n.s.}$, $CFI = .98$, $RMSEA = .05$ [.00; .10], $SRMR = .06$) akzeptable Modellfits auf. Es kann somit davon ausgegangen werden, dass die beiden Testformen A und B sowie die drei Materialarten figural, numerisch und verbal das selbe Konstrukt und dieses jeweils gleich gut messen. Die Testformen A und B des TEMEKKO können als essentiell-parallele Testformen angesehen werden. Darüber hinaus kann abgeleitet werden, dass die

Testleistungen der Untertests lediglich durch eine experimentelle Variation der Aufgaben auf das unterschiedliche Material zurückgeführt werden können.

Die alternative Überprüfung der verwendeten Materialien (figural, numerisch und verbal) anhand von Kovarianzanalysen mit Alter und Geschlecht als Kovariate ergab für den 3-fach Vergleich sowie für alle drei paarweisen Vergleiche zunächst keine statistisch signifikanten Unterschiede. Die Nullhypothese (Gleichheit der Materialien) konnte vorläufig beibehalten werden. Zur Untersuchung, wie groß die Mittelwertsdifferenzen sein müssten, damit von einem praktisch bedeutsamen Unterschied gesprochen werden könnte, wurde vertiefend berechnet. Unter der Annahme des kleinst möglichen Effekts ($\eta^2 = .01$) und einer Teststärke von 95 %, die wiederum durch die Bonferroni-Holm Korrektur nach der Höhe der Effekte angepasst wurde, wurden die hypothetischen kritischen F -Werte ermittelt und mit den zuvor festgestellten empirischen F -Werten abgeglichen. Lediglich eine Paarung (figural : verbal) ergab unter diesen Bedingungen einen statistisch signifikanten Unterschied. Gleichwohl lässt sich mit der empirisch ermittelten Effektgröße von $\eta^2 = .001$ keineswegs von einem praktisch bedeutsamen Unterschied sprechen. Dennoch wurden der Ordnung halber unterschiedliche Normen für die Materialarten bereit gestellt. Die Ergebnisse der Berechnungen sind in nachfolgender Tabelle I.4 dargestellt.

Tabelle I.4: Prüfung der Materialien auf Gleichheit

Material	F_{emp}	p	η^2	power'	F_{krit}
figural : numerisch	0.000	.989	<.001	.950	0.415
numerisch : verbal	0.189	.664	.001	.983	0.440
figural : verbal	0.134	.715	.001	.975	0.107
figural : numerisch : verbal	0.096	.900	.001	.950	1.093

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert

Hinsichtlich der untersuchten Testformen A und B auf deren Unterschiedlichkeit anhand von Kovarianzanalysen nach gleichem Schema zeigte sich, dass die empirisch ermittelte Mittelwertsdifferenz der figuralen Formen A und B statistische Signifikanz erreichte. Naturgemäß erlangte dieser Unterschied bei Abgleich mit dem hypothetisch ermittelten kritischen F -Wert zur Prüfung auf Gleichheit nach den strengeren Kriterien auch statistische Signifikanz. Die Effektstärke ($\eta^2 = .07$) wird als mittlere Größe eingestuft und muss vor dem Hintergrund der strengen Prüfkriterien als praktisch bedeutsam bewertet werden. Tabelle I.5 zeigt die Ergebnisse dieser Analysen. An dieser Stelle sei nicht unerwähnt, dass die Leistungswerte der Testform A im figuralen Teil, der chronologisch als erstes bearbeitet wurde, nominell niedriger lagen, als alle anderen Testteile aller anderen Materialarten, insbesondere als der figuralen Testform B. Aufgrund des fortgeschrittenen Alters der Probanden kann spekuliert werden, ob das Sich-Einfinden in die Aufgabenstellung und in den Bearbeitungsmodus erst eine gewisse Eingewöhnungsdauer erforderten, was letztlich zu geringeren Leistungen in dem als erstes bearbeiteten Testteil führte und möglicherweise die Ursache für die statistisch und praktisch bedeutsamen Mittelwertsdifferenzen der beiden Testformen A und B des figuralen Materials war.

Tabelle I.5: Prüfung der Testformen auf Gleichheit

Testformen	F_{emp}	p	η^2	power'	F_{krit}
A : B (figural)	10.173	.002	.071	.983	0.500
A : B (numerisch)	1.347	.248	.010	.975	3.188
A : B (verbal)	0.774	.381	.006	.950	4.530

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert.

4.5 Reliabilität

Tabelle I.6 stellt die Reliabilitätsschätzungen der in dieser Studie verrechneten Tests anhand ihrer Tempowerte dar.

Tabelle I.6: Reliabilitätsmaße der verwendeten Verfahren

Test	<i>N</i>	<i>r_{tt}</i>
TEMEKKO figural	138	.79 ^a
TEMEKKO numerisch	138	.89 ^a
TEMEKKO verbal	138	.90 ^a
TEMEKKO gesamt	138	.93 ^a
Test d2 ^f	3.176	.97 ^b
TAP, geteilte Aufmerksamkeit ^f	537	
TAP, Go/Nogo ^f	214	
TAP, visuelles Scanning ^f	163	
TAP, Arbeitsgedächtnis ^f	160	
TAP, Reaktionswechsel ^f	588	

^a Paralleltestreliabilität

^b Cronbach α

^c Guttman Split-Half Koeffizient

^f Reliabilitätsschätzungen aus Testmanual

4.6 Validität

Die Inspektion der Korrelationsmatrix aller dieser Untersuchung zugrunde liegenden Testleistungen ließ erste Zusammenhänge des TEMEKKO mit anderen Analysevariablen erkennen. Insbesondere die hohen und signifikanten Korrelationen mit dem Test d2 und Reaktionswechsel (TAP) deuteten auf die Nähe der Konzepte hin. Dieser Hinweis wurde in den nachfolgenden Berechnungen näher analysiert. Tabelle I.7 zeigt zunächst die Korrelationsmatrix der Analysevariablen.

Tabelle I.7: Korrelationsmatrix der Analysevariablen

<i>r</i>	Te fig	Te num	Te verb	Te g	d2	ga	rw	ag	vs	gn	al
Te fig	1	.46**	.43**	.75**	.35**	.32**	.34**	.15	.22**	.16	.07
Te num	.46**	1	.64**	.88**	.54**	.20*	.36**	.29**	.23**	.22*	.11
Te verb	.43**	.64**	1	.84**	.50**	.20*	.42**	.28**	.27**	.15	.20*
Te g	.75**	.88**	.84**	1	.57**	.29**	.46**	.30**	.29**	.21*	.15
d2	.35**	.54**	.50**	.57**	1	.18*	.41**	.22**	.19*	.26**	.18*
ga	.32**	.20*	.20*	.29**	.18*	1	.13	.15	.17	.28**	.24**
rw	.34**	.36**	.42**	.45**	.41**	.13	1	.32**	.43**	.16	.23**
ag	.15	.29**	.28**	.30**	.22**	.15	.32**	1	.18*	.24**	.24**
vs	.22**	.23**	.27**	.29**	.19*	.17	.43**	.18*	1	.21*	.34**
gn	.16	.22*	.15	.22*	.26**	.28**	.16	.24**	.21*	1	.42**
al	.07	.11	.20*	.15	.18*	.24**	.23**	.24**	.34**	.42**	1

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO gesamt; d2 = Test d2; ga = geteilte Aufmerksamkeit, Quadrate, TAP; rw = Reaktionswechsel, TAP; ag = Arbeitsgedächtnis, TAP; vs = Visuelles Scanning, TAP; gn = Go/Nogo, TAP; al = Alertness, TAP. *r* = bivariate Korrelation, * = $p < .05$, ** = $p < .01$.

Zur Untersuchung der von Moosbrugger, Goldhammer und Schweizer (2006) postulierten Struktur von Aufmerksamkeitskomponenten wurde ein Messmodell (vgl. 4.1.2. "Prozedere der Datenbearbeitung" im Manual) mit den beiden Komponenten *Perceptual Attention* und *Executive Attention* generiert. Der Faktor *Perceptual Attention* stellte die gemeinsame Varianz (i.S. eines g-Faktors) aller verwendeten Verfahren dar, wohingegen *Executive Attention* als Gruppenfaktor die Residualvarianz der Tests abbildete, die exekutive Kontrolle erforderten. Die Interpretation des Faktors *Executive Attention* nach Moosbrugger et al. (2006) ist eng verwoben mit der Auffassung von Konzentration, wie sie dem TEMEKKO zu Grunde liegt. Aus funktionaler Sicht lässt sich die Fähigkeit, die *Executive Attention* darstellt, als Steuerung höherer kognitiver Fähigkeiten verstehen, i.e. als exekutive Kontrolle.

Die Prüfung des Modells zeigte mit ($\chi^2[14] = 23.03$, $p = \text{n.s.}$, $CFI = .95$, $RMSEA = .07$ [.00; .12], $SRMR = .05$) einen akzeptablen Modellfit. Der TEMEKKO war stärkster Indikator ($\lambda = .79$) für *Executive Attention*, gefolgt vom Test d2 ($\lambda = .60$), der als Test für

Aufmerksamkeit und Konzentration einfluss. Reaktionswechsel (TAP) als Maß für die Umstellfähigkeit (shifting) stellte mit $\lambda = .49$ ebenso einen wichtigen Indikator für Executive Attention dar. Die Ladungen der Indikatoren TEMEKKO, Test d2 und Reaktionswechsel (TAP) waren auf dem Faktor Executive Attention auf dem 1 % Niveau signifikant und in ihrer Höhe deutlich über ihren jeweiligen Ladungen auf dem Faktor Perceptual Attention. Nichtsdestotrotz erreichten diese auch statistische Signifikanz; i.e. $\lambda = .27$, $p = .01$ für den TEMEKKO, $\lambda = .28$, $p = < .01$ für den Test d2 und $\lambda = .36$, $p < .001$ für Reaktionswechsel.

Interessanterweise deuteten die Ladungen des Tests Geteilte Aufmerksamkeit (TAP) mit $\lambda = .35$, ($p = < .001$) auf Perceptual Attention und $\lambda = .19$ (n.s.) auf Executive Attention darauf hin, dass zur Erfüllung der Aufgabe dem wahrnehmungsgebundenen Aspekt der Aufmerksamkeit eine wichtigere Rolle zukam. Bemerkenswert deshalb, weil von Tests dieser Art, die sich des Dual-Task Paradigmas bedienen, erwartet wird, dass sie Interferenzen verursachen. Demnach wird unterstellt, dass sie exekutive Kontrolle erfordern (Engle, Tuholski, Laughlin & Conway, 1999). Nichtsdestotrotz war dieses Ergebnis konsistent mit dem Befund bei Moosbrugger, Goldhammer und Schweizer (2006), wo geteilte Aufmerksamkeit (TAP) mit $\lambda = .59$ ebenso höher auf Perceptual Attention lud als auf Executive Attention ($\lambda = .24$).

Die übrigen Indikatoren für Executive Attention, die TAP Untertests Arbeitsgedächtnis und visuelles Scanning, luden mit $\lambda = .25$ bzw. $\lambda = .23$ jeweils auf dem 5 % Niveau signifikant auf diesem Faktor. Ihre Ladungen auf Perceptual Attention waren mit $\lambda = .36$ bzw. $\lambda = .45$ jeweils auf dem 1 % Niveau signifikant und deutlich höher als auf Executive Attention. Für diese beiden Verfahren schienen - ähnlich wie für geteilte Aufmerksamkeit - die Erkennung und Selektion von Reizen bedeutsamere Anteile an der Gesamtleistung

beizutragen als die interne Kontrolle innerer Prozesse. Die stärksten Indikatoren für Perceptual Attention waren die beiden TAP Untertests Alertness ($\lambda = .71$, $p < .001$) und focussed attention ($\lambda = .59$, $p < .001$) (über Go/Nogo operationalisiert), was sich aus der konzeptionellen Aufgabenstellung der Tests ableiten ließ. Ein einfaches Reaktionszeitparadigma (Alertness) erfordert keine Änderung oder Neuordnung eines Aufgabensets i.S. der exekutiven Kontrolle der Theory of Visual Attention (EC-TVA). Fokussierte Aufmerksamkeit (Go/Nogo) macht zwar eine Stimulusselektion für die entsprechend richtige Reaktionsauswahl erforderlich, dies bedarf aber nicht der Übertragung neuer Kontrollparameter gemäß der EC-TVA. Die Ergebnisse der Modellprüfung dieser Untersuchung sind in nachfolgender Abbildung I. 1 dargestellt.

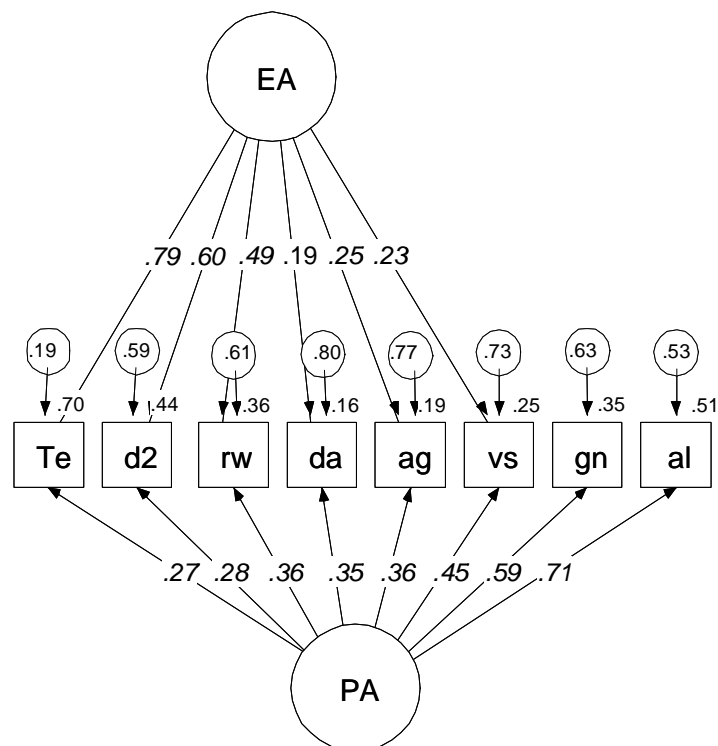


Abbildung I. 1: Modell in Anlehnung an Moosbrugger, Goldhammer und Schweizer (2006)

Anmerkungen: Te = TEMEKKO; d2 = Test d2; rw = Reaktionswechsel, TAP; ga = geteilte Aufmerksamkeit, Quadrate, TAP; ag = Arbeitsgedächtnis, TAP; vs = Visuelles Scanning, TAP; gn = Go/Nogo, TAP; al = Alertness, TAP. EA = Executive Attention; PA = Perceptual Attention.

Aus der Perspektive einer linearen multiplen Regressionsanalyse mit der Testleistung des TEMEKKO als abhängige Variable ließen sich 40 % des erklärbaren Varianzanteils des

Kriteriums durch die verwendeten Analysevariablen vorhersagen. Den deutlich höchsten einzigartigen Varianzanteil erklärte der Test d2 (14 %). Die beiden TAP Untertests Reaktionswechsel und Geteilte Aufmerksamkeit konnten über den Test d2 hinaus noch jeweils einzigartige Anteile von 3 % am TEMEKKO vorhersagen, wobei Reaktionswechsel den weitaus größeren gemeinsamen Varianzanteil mit dem TEMEKKO für sich verzeichnete ($r^2_{simple} = .21$). Dieser Varianzanteil enthielt offenkundig auch Anteile an der gemeinsamen Varianz eines anderen Prädiktors (Test d2), der sich in dessen individueller Vorhersage des Kriteriums zeigte. Die übrigen Variablen enthielten nahezu keine spezifischen Anteile (quadrierte Semipartialkorrelationen), über die bereits genannten Maße hinaus, die entscheidend zur Erklärung des TEMEKKO beitragen konnten. Nichtsdestotrotz waren die gemeinsamen Anteile (quadrierte bivariate Korrelationen) am Kriterium überwiegend hoch signifikant. Lediglich Alertness (TAP) war erwartungsgemäß sowohl hinsichtlich einer zu Grunde liegenden Gemeinsamkeit unbedeutend als auch zur Vorhersage des TEMEKKO nicht geeignet. Tabelle I.8 zeigt die Ergebnisse der Regressionsanalyse.

Tabelle I.8: Regressionsanalyse mit TEMEKKO als Kriterium

	β	r^2_{simple}	$r^2_{semi-partial}$
d2	.42 ***	.32 ***	.14 ***
rw	.20 *	.21 ***	.03 *
ga	.18 *	.09 ***	.03 *
ag	.11	.09 ***	.01
vs	.10	.09 ***	.01
gn	.01	.05 *	< .001
al	-.07	.02	< .01

AV = TEMEKKO, $R^2 = .40$

Anmerkungen: d2 = Test d2; rw = Reaktionswechsel, TAP; ga = geteilte Aufmerksamkeit, Quadrate, TAP; ag = Arbeitsgedächtnis, TAP; vs = Visuelles Scanning, TAP; gn = Go/Nogo, TAP; al = Alertness, TAP. β = standardisierter Regressionskoeffizient, r^2_{simple} = quadrierte bivariate Korrelation, $r^2_{semi-partial}$ = quadrierte Semipartialkorrelation, AV = abhängige Variable; R^2 = korrigierter Determinationskoeffizient; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Zusammenfassung der Validitätsschätzungen

Das zu Grunde gelegte Strukturmodell der Aufmerksamkeit (Moosbrugger, Goldhammer & Schweizer, 2006) konnte repliziert werden. Dies nicht nur hinsichtlich der beiden postulierten Komponenten Executive und Perceptual Attention, sondern auch hinsichtlich des Konzepts, das Executive Attention markierte. Moosbrugger et al. (2006) verwendeten den Frankfurter Adaptiven Konzentrationstest (FAKT) (Moosbrugger & Goldhammer, 2005), der Unterscheidungsaufgaben mit einfachen grafischen Items und formal ähnlichen Distraktoren beinhaltet. Moosbrugger et al. (2006) schreiben dem Verfahren exekutive Aufmerksamkeitsanforderungen zu, weil zwischen Aufgabensets anhand zweierlei Stimulusdimensionen gewechselt werden muss. In der vorliegenden Studie erwies sich ebenso jenes Verfahren als Markertest für Executive Attention, dem der bedeutendste Anteil exekutiver Aufmerksamkeit zugeschrieben wurde, der TEMEKKO ($\lambda = .79$). Die beiden ebenso maßgeblichen Indikatoren für diesen Faktor - die Tests d2 und Reaktionswechsel (TAP) - zeigten Ladungen i.H.v. $\lambda = .60$ und $\lambda = .49$. Für diese drei Verfahren wurde durch deutlich höhere Zuordnungen zum exekutiven Faktor ersichtlich, dass sie mehr Anforderungen an die Konzentration bzw. exekutive Kontrolle stellten, als an wahrnehmungsgebundene Aspekte der Aufmerksamkeit. Dies galt insbesondere für den TEMEKKO.

Der TEMEKKO markierte den Faktor Executive Attention und wurde somit als adäquater Repräsentant der Konzentration bzw. exekutiven Kontrolle auf funktionaler Ebene verstanden. Auch dem Test d2 kann aus Sicht der Aufgabenstellung unterstellt werden, dass er hauptsächlich die Steuerung/Kontrolle von untergeordneten Prozessen bei der Stimulusselektion sowie der Reaktionsauswahl (response selection) erfordert. Die Fähigkeit zum Wechsel des Aufmerksamkeitsfokus, wie sie vom Test Reaktionswechsel

gefordert wird, wird als spezifische oder allgemeine Form der kognitiven Flexibilität angesehen.

Die wahrnehmungsgebundene latente Variable Perceptual Attention wurde durch Alertness ($\lambda = .71$) markiert, gefolgt von fokussierter Aufmerksamkeit (Go/Nogo) ($\lambda = .59$). Diese Verfahren basieren auf perzeptueller Kategorisierung bei der visuellen Erkennung und Selektion von Reizen und stellen gemäß EC-TVA keine Anforderungen an die exekutive Kontrolle.

Zur Vorhersage des erklärbaren Varianzanteils des TEMEKKO (40 %) eignete sich in erster Linie der Test d2 mit dem weitaus größten einzigartigen Anteil (14 %) an der Prädiktion. Umstellfähigkeit (Reaktionswechsel, TAP) erwies sich als zweitbesten Prädiktor ($\beta = .20$). Zwar teilte er einen maßgeblichen gemeinsamen Varianzanteil mit dem Kriterium (21 %), dieser enthielt allerdings nur eine relativ geringe einzigartige Vorhersagekraft (3 %). Einen ähnlich großen einzigartigen Beitrag an der Erklärung des TEMEKKO zeigte geteilte Aufmerksamkeit (TAP) ($\beta = .18$), obwohl der gemeinsame Varianzanteil mit dem Kriterium (9 %) deutlich geringer ausfiel.

Validierungsstudie 2

Der Datensatz, der als Validierungsstudie 2 diente, stammt aus einer Untersuchung einer Diplomandin der Philipps-Universität Marburg, Fachbereich Psychologie.

1 Testdurchführung

Die Daten wurden in der Zeit von Oktober bis Dezember 2002 im diagnostischen Labor im Institutsgebäude des Fachbereichs Psychologie der Philipps-Universität Marburg erhoben. Die Testung wurde hinsichtlich der technischen Einrichtungen und des Prozederes nahezu voll standardisiert. Die Untersuchung fand grundsätzlich als Gruppentestung statt, bei der jeweils fünf Personen gleichzeitig die Tests an Computern durchführten. Eine Person wurde aufgrund ihrer unentschuldigten Abwesenheit zum Testtermin einer Einzeluntersuchung unterzogen.

Zu Beginn der Testung wurden die Versuchspersonen darüber informiert, dass die Rückmeldung ihrer Testergebnisse per E-Mail oder anonym via verschlossenem und ausschließlich mit der Teilnehmernummer versehenem Kuvert, das von den Probanden abgeholt werden konnte, erfolgen könne. Im Anschluss daran wurden die anonymisierten Kopien der Abiturzeugnisse eingesammelt und ein Fragebogen zur Erfassung der Personendaten von den Versuchspersonen ausgefüllt. Ein Überblick über den Ablauf der Testung durch die Versuchsleiterin schloss den Eingangsteil der Untersuchung.

Daran anschließend erfolgte die eigentliche Testung der Versuchspersonen. Die Reihenfolge der Testdarbietung wurde bewusst aus folgenden Gründen gewählt: Zu Beginn wurde der TEMEKKO dargeboten, um die Konzentrationsleistung der Probanden möglichst rein, ohne Ermüdungseinflüsse von zuvor durchgeführten Tests, zu erfassen. Ebenso sollten die Versuchspersonen bei der Beantwortung der Aufgaben der Testbatterie zur Aufmerksamkeitsprüfung (TAP) noch so wenig wie möglich ermüdet sein; darum wurden diese Tests direkt im Anschluss an den TEMEKKO gestellt. Als letztes Verfahren der Untersuchung wurde der Intelligenztest I-S-T 2000 R durchgeführt. Es war davon auszugehen, dass die Probanden hinsichtlich der Rückmeldung ihrer Testergebnisse am meisten an ihren Intelligenzleistungen Interesse hatten. So konnte sichergestellt werden, dass ihre Motivation bis zum letzten Verfahren der Testbatterie aufrecht erhalten werden konnte.

Technische Schwierigkeiten (Computer Absturz) bei der Durchführung der TAP-Untertests, Geteilte Aufmerksamkeit und Alertness führten bei einigen wenigen Probanden dazu, dass die Computer während der Testung neu gestartet werden mussten. Fand der Absturz zu Beginn des Untertests statt, wurde dieser wiederholt. Musste der PC eher gegen Ende des Tests neu gestartet werden, wurde dieser Untertest für die jeweilige Versuchsperson abgebrochen.

2 Stichprobe

Der Datensatz umfasste 144 Personen im Alter von 19 bis 39 Jahren ($M = 23.7$ Jahre; $SD = 3.37$ Jahre). Das mittlere Alter der 78 weiblichen Versuchspersonen (54.2 %) betrug

23.2 Jahre ($SD = 3.57$ Jahre), das der 66 männlichen Probanden 24.2. Jahre ($SD = 3.05$ Jahre).

Bei drei Probanden aus dieser Stichprobe handelte es sich um wissenschaftliche Mitarbeiter der Universität. Die übrigen 141 Personen waren ausschließlich Studierende. Die Zugehörigkeit ihrer Studienfächer verteilte sich wie folgt: Humanmedizin ($N = 21$), Jura und Psychologie (jeweils $N = 12$), Biologie ($N = 11$), BWL ($N = 9$). Die übrigen 76 Studierenden hatten 29 unterschiedliche Studiengänge belegt. Durchschnittlich lag die Semesterzahl der Versuchspersonen bei 5.6 ($SD = 4.0$).

Die Teilnehmer an dieser Untersuchung wurden von der Diplomandin in erster Linie über einen Link auf der Homepage der Philipps-Universität Marburg angeworben. Zusätzlich wurden im Klinikum der Universität Marburg sowie in der Studentenschaft anderer Fachbereiche Versuchspersonen über Plakatinformationen rekrutiert. Die Teilnahmevoraussetzungen der Interessenten umfassten: Alter zwischen 18 und 35 Jahre, sehr gute Deutschkenntnisse und nicht sehbehindert. Darüber hinaus sollten die Teilnehmer keine Erfahrungen mit den eingesetzten Verfahren gesammelt haben und sie sollten ihre anonymisierten Abiturzeugnisse beibringen.

3 Versuchsmaterial

Die Testverfahren, die die Diplomandin für ihre Untersuchung verwendet hatte, umfassten mehr Verfahren, als in der vorliegenden Studie verwendet wurden (vgl. Tabelle II.1). Für die statistischen Analysen zur Validierung des TEMEKKO wurden Testvariablen zu den überlappenden bzw. abzugrenzenden Konstrukten ausgewählt. Die Variablen des

TEMEKKO wurden mutmaßlich als Indikatoren für Konzentration verwendet. Drei Skalen des I-S-T 2000 R (verbale Intelligenz, numerische Intelligenz, figurale Intelligenz) flossen als Indikatoren für Intelligenz/Reasoning in die Berechnungen mit ein, drei Untertests der Testbatterie zur Aufmerksamkeitsprüfung (TAP), i.e. geteilte Aufmerksamkeit, Go/Nogo und Alertness für Aufmerksamkeit sowie drei Tests der Arbeitsgedächtnis-Testbatterie von Oberauer, Süß, Schulze, Wilhelm und Wittmann (2000), i.e. Reading Span, Figurale Koordination sowie Switching Numerisch für die Arbeitsgedächtniskapazität.

In Tabelle II.1 ist der genaue zeitliche Ablauf der Testung dargestellt. Die Testverfahren EQ-iBA (Schmidt-Atzert & Bühner, unveröffentlichtes Verfahren) sowie MultiFlux (Kröner, 2001) werden der Vollständigkeit halber gelistet. Die Daten aus diesen beiden Verfahren wurden für andere Forschungszwecke erhoben. Sie waren für die vorliegende Validierungsstudie nicht relevant

Tabelle II.1: Ablaufplan der Testung

Testverfahren	Untertest	Dauer	Art der Bearbeitung
Erfassung Personendaten		5 min.	Paper-Pencil
TEMEKKO		20 min.	Paper-Pencil
Testbatterie zur Aufmerksamkeitsprüfung (TAP)	Alertness	40 min.	PC-gestützt
	Go/Nogo (Bedingung 2)	(für alle Subtests)	PC-gestützt
	Geteilte Aufmerksamkeit (Serie 4, Test C, 60 Trials)		PC-gestützt

Tabelle II.1 wird fortgesetzt

Arbeitsgedächtnistests	Switching Numerisch Reading Span Figurale Koordination	40 min. (für alle Subtests)	PC-gestützt mit Paper-Pencil- Komponente (für alle Subtests)
EQ-iBa		35 min.	Paper-Pencil
<i>Pause</i>		10 min.	
MultiFlux		45 min.	PC-gestützt
I-S-T 2000 R		80 min.	PC-gestützt

3.1 TEMEKKO

Für die detaillierte Beschreibung des TEMEKKO sei auf Abschnitt 3 (Entwicklung des TEMEKKO) des Manuals verwiesen.

3.2 Testbatterie zur Aufmerksamkeitsprüfung (TAP)

Für die Beschreibung des Verfahrens sowie zu den Untertests Alertness, geteilte Aufmerksamkeit und Go/Nogo, die in dieser Studie als Verfahren verwendet wurden, sei auf Abschnitt 3.3 (Testbatterie zur Aufmerksamkeitsprüfung) und die zugehörigen Unterpunkte 3.3.1 (Alertness), 3.3.2 (geteilte Aufmerksamkeit) und 3.3.3 (Go/Nogo) in Anhang I (Validierungsstudie 1) verwiesen.

3.3 Arbeitsgedächtnis-Testbatterie (2000)

Eine umfangreiche Testbatterie zur Erfassung der Arbeitsgedächtniskapazität (Oberauer Süß, Schulze, Wilhelm & Wittmann, 2000) wurde in deren Studie faktorenanalytisch untersucht und brachte drei Faktoren hervor. Inhaltlich wurden diese als (1) simultanes Speichern und Verarbeiten bzw. Koordination von verbal-numerischem Material, (2) simultanes Speichern und Verarbeiten bzw. Koordination von räumlich-visuellem Material und (3) Supervision und Speed (Bearbeitungsgeschwindigkeit) interpretiert. Aus dieser Testbatterie (insgesamt 23 Subtests) wurden für die vorliegende Validierungsstudie drei Untertests ausgewählt, die als Markertests die jeweiligen Komponenten besonders gut repräsentierten. Die ausgewählten Tests zeigten hohe (wenn auch nicht die höchsten) Ladungen auf den gefundenen Faktoren. Aufgrund ihrer häufigen Verwendung in anderen Untersuchungen wurden sie jenen Tests mit der höchsten Ladung auf dem jeweiligen Faktor vorgezogen. So wurde der Test Reading Span als Markertest für den Faktor simultanes Speichern und Verarbeiten bzw. Koordination von verbal-numerischem Material mit einer Ladung von $a = .71$ ($a = .85^1$) ausgewählt. Ebenso diente der Test figurale Koordination mit einer Ladung von $a = .76$ ($a = .77^1$) als Repräsentant für den Faktor simultanes Speichern und Verarbeiten bzw. Koordination von räumlich-visuellem Material. Für den Faktor Supervision und Speed (Bearbeitungsgeschwindigkeit) galt der Test Switching numerisch mit einer Ladung von $a = .74$ ($a = .80^1$) als Indikator.

Die Darbietung der Aufgaben erfolgte vollständig computergestützt, wohingegen die Instruktionen zu den Aufgaben lediglich in Papierform vorlagen und vom Versuchsleiter

¹ höchste Ladung eines anderen Tests auf dem jeweiligen Faktor

zusätzlich vorgelesen wurden. Teilweise wurden die Antworten der Versuchspersonen computergestützt erfasst, teilweise handschriftlich in Antworthefte eingetragen.

3.3.1 Switching numerisch

Dieser Test gilt als Indikator für den Faktor Supervision/Speed (Bearbeitungsgeschwindigkeit). Bei dieser Aufgabe muss zwischen zwei einfachen Aufgaben bei gleichem Stimulusmaterial hin- und hergewechselt werden. Eine unterschiedliche Anzahl von gleichen Ziffern wird am Bildschirm in zufälliger Anordnung präsentiert. Die Versuchspersonen sollen so schnell als möglich jeweils abwechselnd den Wert der dargebotenen Ziffer und bei Erscheinen des nächsten Reizes (der nächsten Gruppe von Ziffern) die Anzahl der dargebotenen Ziffern eintippen. Die aktuell geltende Instruktion wird jeweils am oberen Rand des Bildschirms angezeigt („Anzahl eingeben“ oder „Ziffer eingeben“). Die einzugebenden Zahlen sind immer einstellig (1 bis 9), wobei der Ziffernwert und die Anzahl der Ziffern in keiner Aufgabe deckungsgleich sind. Die Darbietung der nächsten Aufgabe erfolgt jeweils 400 ms nach Eintippen der Antwort. Es werden 10 Übungs- und 30 Testaufgaben dargeboten. Als Leistungsmaß diente die mittlere Reaktionszeit der richtig beantworteten Items.

3.3.2 Reading Span

Für den Faktor Simultanes Speichern und Verarbeiten bzw. Koordination von verbal-numerischem Material wird Reading Span als Markertest verwendet. Der Versuchsperson wird eine bestimmte Anzahl von kurzen Sätzen nacheinander dargeboten. Jeder Satz

erscheint drei Sekunden lang auf dem Bildschirm, gefolgt von einer Sekunde Interstimulusintervall bis zum nächsten Satz. In diesen vier Sekunden muss die Versuchsperson den Inhalt des Satzes nach wahr („J“) oder falsch („N“) durch Tastendruck bewerten. Die Sätze umfassen triviale Aussagen, sodass möglichst wenig Vorwissen erforderlich ist. Über die Beurteilung des Satzinhaltes hinaus muss die Versuchsperson das jeweils letzte Wort eines Satzes memorieren und nach einem Set von Sätzen (zwischen 3 und 7 Sätzen) die gemerkten Wörter in der Reihenfolge der Darbietung auf ein Antwortblatt schreiben. Diese Wörter sind Substantive mit relativ hoher Worthäufigkeit, zwischen denen keine offensichtlichen semantischen Beziehungen bestehen. Der Test beinhaltet zwei Übungs- und 15 Testaufgaben. Als Leistungsmaß diente die mittlere Anzahl der an der richtigen Position erinnerten Wörter.

3.3.3 Figurale Koordination

Der Repräsentant des Faktors Simultanes Speichern und Verarbeiten bzw. Koordination von räumlich-visuellem Material ist der Test Figurale Koordination. Den Versuchspersonen wird eine 10 x 10 Felder Matrix dargeboten, in der nacheinander für jeweils eine Sekunde zwei oder mehr (bis zu sechs) Punkte erscheinen. Das so entstandene Punktemuster ist zu memorieren. Nach einer entsprechenden Aufforderung, die auf dem Bildschirm erscheint, ist das erinnerte Punktemuster auf einem Antwortbogen einzutragen. Hierbei ist nicht die absolute Position der Punkte maßgeblich, sondern ihre relative Anordnung zueinander. Als zusätzliche Aufgabe gilt es, vor der Wiedergabe des Punktemusters zu entscheiden, ob das Muster entlang einer imaginierten Vertikalen symmetrisch ist oder nicht. In Abhängigkeit davon wird das erinnerte Punktemuster in die linke („symmetrisch“) oder rechte („nicht symmetrisch“) Matrix auf dem Antwortbogen

eingetragen. Nach zwei Übungsaufgaben werden 15 Testaufgaben dargeboten. Als Leistungsmaß diente die mittlere Anzahl der richtig positionierten Punkte.

3.4 Intelligenz-Struktur-Test 2000, revidierte Fassung (I-S-T 2000 R)

(Amthauer, Brocke, Liepmann & Beauducel, 2001). Das Verfahren erfasst eine möglichst große Bandbreite der fluiden und kristallisierten Intelligenz sowie schlussfolgernden Denkens, wie sie als Konzepte von Cattell (1963, 1987) und in weiteren Forschungen (z.B. Cattell & Horn, 1978; Horn, 1988) entwickelt wurden. Die Grundmodul-Kurzform - ohne Merkfähigkeitsaufgaben - besteht aus neun Untertests, die drei inhaltsbezogene Fähigkeiten (verbal, numerisch, figural) erfassen. Jeder Subtest beinhaltet 20 Aufgaben. Die Rohwerte der Tests Satzergänzung, Analogien und Gemeinsamkeiten bildeten die Skala verbale Intelligenz. Numerische Intelligenz setzte sich aus den Untertests Rechenaufgaben, Zahlenreihen und Rechenzeichen zusammen. Die Skala figurale Intelligenz wurde durch die Tests Figurenauswahl, Würfelaufgaben und Matrizen gebildet. Aus diesen Skalen wurde der Gesamtwert des schlussfolgernden Denkens durch Summation berechnet. Das Verfahren wird den Versuchspersonen vollständig computergestützt dargeboten. Die Reihenfolge der Bearbeitung der einzelnen Untertests ist durch die Programmierung des Verfahrens festgelegt und entspricht der Darstellung in Tabelle II.2. Für jeden Test ist die Bearbeitungszeit begrenzt. Gleichwohl hat der Proband die Möglichkeit, innerhalb dieses Zeitrahmens einzelne Aufgaben zu überspringen und zu einem späteren Zeitpunkt darauf zurück zu kommen. Bei allen Subtests dieser Grundmodul-Kurzform werden jeweils 20 Aufgaben dargeboten. Die Auswertung der einzelnen Subtests erfolgte über das Testprogramm und konnte direkt in eine SPSS Datei überführt werden.

Tabelle II.2: Übersicht über die Untertests und Skalen der Grundmodul-Kurzform des I-S-T 2000 R in der Reihenfolge der Testbearbeitung (nach Schmidt-Atzert, 2002)

übergeord. Faktor	Primärfaktoren	Subtests ^a	Aufgabenerläuterung (teilw. mit Beispielitems)	
Schluss- folgerndes Denken	V	Verbale Intelligenz	SE	Satzergänzung (6) Sätze sind um ein Wort zu ergänzen: Schuhe haben / bestehen immer aus ? a) Leder b) Schnürsenkel c) Sohlen d) Fußbett e) Absätze
			AN	Analogien (7) Gleichungen der Wörter sind zu ergänzen: klein : groß = kurz : ? a) breit b) schmal c) lang d) ausgedehnt e) weit
			GE	Gemeinsamkeit n (8) Aus sechs Begriffen sind zwei mit dem gleichen Oberbegriff zu finden: a) Café b) Behausung c) Straße d) Büro e) Garten f) Quartier
	N	Numerische Intelligenz	RE	Rechenaufgaben (10) Aufgabenlösen, mit Hilfe der vier Grundrechenarten sowie weiteren, wie z.B. Bruchrechnen, Potenzieren $50 - 30 = A$
			ZR	Zahlenreihen (10) Zahlenfolgen um eine Zahl fortsetzen 2 5 8 11 14 17 20 ?
			RZ	Rechenzeichen (10) In Gleichungen sind fehlende Rechenzeichen (+, -, x, :) einzufügen $14 ? 8 ? 3 = 66$
			FA	Figurenauswahl (7) Zerschnittene Figuren sind ganzen zuzuordnen
	F	Figurale Intelligenz	WÜ	Würfelaufgaben (9) Würfel mit verschiedenen Zeichen sind Würfeln mit den gleichen Zeichen in anderen Positionen zuzuordnen
			MA	Matrizen (10) Eine Reihe von Figuren ist nach einer bestimmten Regel um eine Figur zu erweitern

^a Angabe der Bearbeitungszeiten in Minuten in Klammern

3.4.1 Satzergänzung (SE)

Bei diesem Test wird pro Aufgabe ein Satz dargeboten, in dem ein Wort fehlt. Aus fünf Wörtern muss jenes Wort ausgewählt werden, das den Satz richtig vervollständigt.

3.4.2 Analogien (AN)

Jede Aufgabe besteht aus drei Begriffen, von denen die ersten beiden in einer bestimmten Beziehung zueinander stehen. Zunächst muss diese Beziehung richtig identifiziert werden, um dann im zweiten Schritt aus fünf Antworten diejenige auszuwählen, die eine ähnliche Relation zum dritten Begriff aufweist.

3.4.3 Gemeinsamkeiten (GE)

Hier wird eine Gruppe von sechs Wörtern präsentiert, aus der zwei Wörter mit einem gemeinsamen Oberbegriff identifiziert werden sollen.

3.4.4 Rechenaufgaben (RE)

Zur Lösung einer Aufgabe werden Rechenoperationen im Bereich der reellen Zahlen verlangt.

3.4.5 Zahlenreihen (ZR)

Bei diesem Test werden Zahlenreihen vorgegeben, die jeweils nach einer bestimmten Regel aufgebaut sind. Die Aufgabe besteht darin, die Zahl zu finden, die die Reihe jeweils richtig ergänzt.

3.4.6 Rechenzeichen (RZ)

Der Proband bekommt Gleichungen im Bereich der rationalen Zahlen dargeboten, in denen die Rechenzeichen ausgelassen wurden. Die Aufgabe besteht darin, die fehlenden Rechenzeichen der vier Grundrechnungsarten richtig einzusetzen.

3.4.7 Figurenauswahl (FA)

Jede Aufgabe zeigt fünf Stücke einer zerschnittenen Figur. Aus zehn zur Auswahl stehenden Figuren soll jene herausgefunden werden, die sich nach Zusammensetzen der einzelnen Teilstücke ergeben würde.

3.4.8 Würfelaufgaben (WÜ)

Bei diesem Test werden verschiedene Würfel präsentiert, die auf jeder Seite ein unterschiedliches Muster aufweisen. Drei Seiten eines Würfels sind für den Betrachter sichtbar. Aus fünf verschiedenen Würfeln soll jener gefunden werden, bei dem es sich um den dargebotenen Würfel in veränderter Position handelt.

3.4.9 Matrizen (MA)

Jede Aufgabe zeigt eine Anordnung von Figuren, die nach einer bestimmten Regel aufgebaut sind. Aus fünf zur Auswahl stehenden Figuren soll jene ausgewählt werden, die die Anordnung regelkonform ergänzt.

4 Statistische Analysen

4.1 Deskriptiva

Zunächst wird ein Überblick über die Rohdaten der Analysevariablen des Datensatzes gelistet. Der Untertest Alertness (TAP) wurde in nachfolgenden Analysen nicht mit berücksichtigt, weil sich aufgrund der Korrelationsmatrix ergab, dass Alertness keinen substantiellen Zusammenhang mit anderen Analysevariablen aufwies und somit als irrelevant angesehen werden konnte. Der Vollständigkeit halber wird Alertness in den Deskriptiva aber gelistet.

Tabelle II.3: Deskriptive Statistiken der Testvariablen

Tests	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>	<i>r_{age}</i>	<i>r_{gender}</i>
Te fig	144	22.0	74.0	44.7	9.2	0.22	0.47	-.10	-.13
Te num	144	17.0	88.0	46.6	13.2	0.62	0.44	.02	-.25**
Te verb	144	25.0	66.0	43.8	8.1	0.42	0.07	-.00	-.10
Te g	144	78.0	217.0	135.0	27.0	0.54	0.24	-.02	-.20*
sfd	144	73.0	159.0	125.8	18.0	-0.42	-0.23	-.02	-.23**
vi	144	29.0	53.0	42.9	5.2	-0.47	-0.27	.11	-.25**
ni	144	16.0	60.0	45.3	10.3	-0.72	0.02	-.01	-.21*
fi	144	19.0	57.0	37.7	7.5	-0.17	-0.50	-.13	-.09
rs	144	0.7	4.9	3.8	0.7	-0.90	1.68	.08	-.04
fk	144	1.7	3.7	2.9	0.4	-0.14	-0.19	-.09	-.21*
sn	144	1,068.4	2,767.7	1,699.5	308.4	0.89	1.38	.11	.15
ga	143	488.0	835.0	638.0	62.0	0.35	0.42	.03	.15
gn	143	326.5	697.5	486.5	60.5	0.26	0.75	-.02	.28**
al	140	173.0	304.5	222.3	27.3	0.77	0.21	-.12	.16

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO gesamt; sfd = schlussfolgerndes Denken, IST 2000 R; vi = verbale Intelligenz, IST 2000 R; ni = numerische Intelligenz, IST 2000 R; fi = figurale Intelligenz, IST 2000 R; rs = Reading Span, Arbeitsgedächtnis-Testbatterie; fk = figurale Koordination, Arbeitsgedächtnis-Testbatterie; sn = Switching numerisch, Arbeitsgedächtnis-Testbatterie; ga = geteilte Aufmerksamkeit, TAP; gn = Go/Nogo, TAP; al = Alertness, TAP. *r* = bivariate Korrelationen, * = $p < .05$, ** = $p < .01$.

Die optische Prüfung der Verteilungsformen der einzelnen Testvariablen mittels Histogrammen und zugehörigen Normalverteilungskurven sowie die Inspektion der numerischen Kennwerte für Schiefe und Kurtosis erbrachten keine außergewöhnlich schiefen oder gewölbten Auffälligkeiten. Die kreuzweise angefertigten Streudiagramme gaben nach augenscheinlicher Prüfung keinen Anlass, Fälle aus den Auswertungen zu eliminieren. Es lagen zwar einige Werte abseits der „Punktwolke“, allerdings im Trend. Es wurden daher alle Fälle für die weiteren Berechnungen beibehalten.

4.2 Zusammenhang mit Alter und Geschlecht

Nach z-Standardisierung der Variablen und Umpolung der Reaktionszeiten (Multiplikation mit -1) wurden die Zusammenhänge mit Alter und Geschlecht ermittelt. Die Korrelationen der Testvariablen zeigten insbesondere mit Geschlecht substantielle Zusammenhänge. Um den Einfluss von Alter und Geschlecht aus den Analysevariablen zu eliminieren, wurden die z-transformierten und ggfs. umgepolten Testkennwerte als abhängige Variable und Alter und Geschlecht als unabhängige Variablen einer linearen multiplen Regressionsanalyse (Einschluss Methode) unterzogen. Zur Prüfung des Zusammenhangs der Testvariablen und jenen, die um Alter und Geschlecht bereinigt wurden, wurden Streudiagramme für jede einzelne Testvariable (standardisiertes Residuum auf der Abszisse, z-transformierter und ggfs. umgepolter Rohwert auf der Ordinate) in Augenschein genommen. Es zeigte sich durchgängig ein positiver linearer Trend. Nachfolgende Tabelle II.4 zeigt die korrelativen Zusammenhänge der Analysevariablen vor und nach Auspartialisierung der Variablen Alter und Geschlecht. Der Ordnung halber sei erwähnt, dass alle Korrelationen auf dem 1 % Niveau signifikant waren.

Tabelle II.4: Zusammenhang z-transformierte Testvariablen und standardisierte Residuen

z \ Res.	Te fig	Te num	Te verb	Te g	sfd	vi	ni	fi	rs	fk	sn	ga	gn
Te fig	.984												
Te num		.968											
Te verb			.995										
Te g				.979									
sfd					.972								
vi						.966							
ni							.977						
fi								.986					
rs									.997				
fk										.970			
sn											.980		
ga												.988	
gn													.959

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO gesamt; sfd = schlussfolgerndes Denken, IST 2000 R; vi = verbale Intelligenz, IST 2000 R; ni = numerische Intelligenz, IST 2000 R; fi = figurale Intelligenz, IST 2000 R; rs = Reading Span, Arbeitsgedächtnis-Testbatterie; fk = figurale Koordination, Arbeitsgedächtnis-Testbatterie; sn = Switching numerisch, Arbeitsgedächtnis-Testbatterie; ga = geteilte Aufmerksamkeit, TAP; gn = Go/Nogo, TAP.

4.3 Schätzung fehlender Werte

Die Schätzung fehlender Werte wurde mittels EM-Algorithmus ergänzt. Der MCAR-Test (missing completely at random) nach Little ergab mit $\chi^2[25] = 23.90$, n.s., dass die MCAR-Nullhypothese akzeptiert wird. Das Fehlen der Werte war demnach unabhängig von sowohl der Ausprägung der fehlenden Werte, als auch von den Ausprägungen aller anderen Werte im Datensatz (Byrne, 2001). Die grafische Überprüfung der standardisierten Residuen mit fehlenden Werten und jener standardisierten Residuen nach Ergänzung der fehlenden Werte anhand von Streudiagrammen sowie der korrelativen Zusammenhänge fiel erwartungsgemäß zufriedenstellend aus. Es waren lediglich zwei Testvariablen (geteilte Aufmerksamkeit und Go/Nogo der TAP), die jeweils einen fehlenden Wert (0.7 %) bei $N = 144$, aufwiesen.

4.4 Prüfung auf Gleichheit der Testformen und Materialarten

Die Prüfung der beiden Strukturgleichungsmodelle (Modell a) und b) in Abbildung 3 des Abschnitts 4.1.2 (Prozedere der Datenbearbeitung) im Manual) auf Passung der Daten wurde unter Berücksichtigung der Bollen-Stine-Bootstrap-Korrektur vorgenommen, da die Annahme einer multivariaten Normalverteilung verletzt war (Tempowerte: multivariate Kurtosis = 5.99, c.r. = 3.67, $p = < .001$ bzw. Fehler-Prozent-Werte: multivariate Kurtosis = 57.89, c.r. = 35.45, $p = < .001$). Die Angaben der Schiefe und Kurtosis der einzelnen Testvariablen zeigten hingegen akzeptable Werte hinsichtlich der von West, Finch und Curran (1995) postulierten Grenzen von Schiefe < 2 und Kurtosis < 7 für die Tempowerte. Hinsichtlich der Fehler-Prozent-Variablen waren die akzeptablen Grenzen der figural-Fehler-B-Prozente mit 3.55 für die Schiefe und 19.85 für die Kurtosis überschritten.

Die Strukturgleichungsmodelle wiesen sowohl bei der Prüfung auf Passung der Tempowerte (Modell a) ($\chi^2 [16] = 22.40$, $p = \text{n.s.}$, $CFI = .99$, $RMSEA = .05$ [.00; .10], $SRMR = .03$) als auch der Fehler-Prozent-Werte (Modell b) ($\chi^2 [16] = 11.62$, $p = \text{n.s.}$, $CFI = 1.00$, $RMSEA = .00$ [.00; .05], $SRMR = .04$) einen akzeptablen Modellfit auf. Es kann demnach davon ausgegangen werden, dass die beiden Testformen A und B sowie die drei Materialarten figural, numerisch und verbal das selbe Konstrukt und dieses jeweils gleich gut messen. Die Testformen A und B des TEMEKKO können somit als essentiell-parallele Testformen anerkannt werden. Darüber hinaus kann abgeleitet werden, dass die Testleistungen der Untertests lediglich auf eine experimentelle Variation der Aufgaben zurückgeführt werden können.

Bei der Prüfung der Materialarten (figural, numerisch und verbal) auf deren Unterschiede zueinander anhand von Kovarianzanalysen mit Alter und Geschlecht als Kovariate ergaben für den 3-fach Vergleich sowie für alle drei paarweisen Vergleiche zunächst keinen statistisch signifikanten Unterschied der Materialien, d.h. dass die Nullhypothese (Gleichheit der Materialien) beibehalten wurde. Um zu prüfen, wie groß ein Effekt sein müsste, damit von einem praktisch bedeutsamen Unterschied gesprochen werden könnte, wurde vertiefend berechnet. Unter der Annahme des kleinst möglichen Effekts ($\eta^2 = .01$) und einer Bonferroni-Holm Anpassung der Teststärke (von ursprünglich durchwegs 95 %) wurden die hypothetischen kritischen F -Werte ermittelt und mit den zuvor festgestellten empirischen F -Werten abgeglichen. Dieser Vergleich ergab für die Paarungen numerisch : verbal und figural : verbal statistisch wie auch praktisch bedeutsame Unterschiede. Die in diesen Vergleichen erreichten Effektgrößen von jeweils $\eta^2 = .02$ wurden allerdings als gering eingestuft. Nichtsdestotrotz wurden für die Materialarten unterschiedliche Normen des TEMEKKO zur Verfügung gestellt. In Tabelle II. 5 sind die Ergebnisse der Berechnungen dargestellt.

Tabelle II. 5: Prüfung der Materialien auf Gleichheit

Material	F_{emp}	p	η^2	power'	F_{krit}
figural : numerisch	0.051	.821	< .001	.950	1.644
numerisch : verbal	2.880	.092	.020	.975	1.464
figural : verbal	3.103	.080	.022	.983	0.441
figural : numerisch : verbal	1.797	.168	.013	.950	2.155

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert

Hinsichtlich der untersuchten Testformen A und B auf deren Unterschiedlichkeit bzw. Gleichheit anhand von Kovarianzanalysen nach gleichem Schema ließ sich festhalten, dass die empirisch ermittelten Mittelwertsdifferenzen sowohl vor als auch nach dem Prozedere des Abgleichs anhand des kleinstmöglich angenommenen Effektes weder statistische noch

praktische Bedeutung erlangten. Die Testformen A und B konnten in allen Fällen als gleich angesehen werden. Die Ergebnisse hierzu sind in Tabelle II.6 gelistet.

Tabelle II.6: Prüfung der Testformen auf Gleichheit

Testformen	F_{emp}	p	η^2	power'	F_{krit}
A : B (figural)	2.320	.130	.016	.983	3.371
A : B (numerisch)	1.782	.184	.012	.975	14.293
A : B (verbal)	0.287	.593	.002	.950	7.681

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert.

4.5 Reliabilität

Zur Übersicht der in dieser Untersuchung verwendeten Tests zeigt Tabelle II.7 die Reliabilitätsschätzungen ihrer Tempowerte.

Tabelle II.7: Reliabilitätsmaße der verwendeten Verfahren

Test	N	r_{tt}
TEMEKKO figural	144	.90 ^a
TEMEKKO numerisch	144	.96 ^a
TEMEKKO verbal	144	.92 ^a
TEMEKKO gesamt	144	.97 ^a
IST 2000 R, verbale Intelligenz ^d	3.484	.88 ^b
IST 2000 R, numerische Intelligenz ^d	3.484	.95 ^b
IST 2000 R, figurale Intelligenz ^d	3.484	.87 ^b
IST 2000 R, schlussfolgerndes Denken ^d	3.484	.96 ^b
Arbeitsgedächtnis, Reading Span ^e	128	.84 ^b
Arbeitsgedächtnis, Figurale Koordination ^e	128	.75 ^b
Arbeitsgedächtnis, Switching Numerisch ^e	128	.78 ^b
TAP, tonische Alertness ^d	308	~ .99 ^c
TAP, geteilte Aufmerksamkeit ^d	537	.99 ^c
TAP, Go/Nogo ^d	214	~ .99 ^c

^a Paralleltestreliaibilität

^b Cronbach α

^c Guttman Split-Half Koeffizient

^d Reliabilitätsmaße aus Testmanual

^e Reliabilitätsmaße aus Studie Oberauer, Süß, Schulze, Wilhelm und Wittmann (2000)

4.6 Validität

Die in Tabelle II.8 dargestellte Korrelationsmatrix der Analysevariablen wurde in Augenschein genommen und ermöglichte erste Erkenntnisse über Zusammenhänge zwischen den Testvariablen. Besonders auffällig hierbei erschienen die hohen Korrelationen der TEMEKKO Subtests mit numerischer und figuraler Intelligenz aus dem I-S-T 2000 R. Auch die deutlichen Zusammenhänge mit den Arbeitsgedächtnistests Figurale Koordination, Reading Span und Switching Numerisch, gaben erste Hinweise auf Gemeinsamkeiten, die es näher zu untersuchen galt.

Tabelle II.8: Korrelationsmatrix der Analysevariablen

<i>r</i>	Te fig	Te num	Te verb	Te g	vi	ni	fi	ga	gn	sn	fk	rs
Te fig	1	.66**	.62**	.85**	.27**	.40**	.51**	.28**	.25**	.40**	.44**	.35**
Te num	.66**	1	.71**	.93**	.28**	.56**	.55**	.24**	.23**	.48**	.47**	.39**
Te verb	.62**	.71**	1	.86**	.25**	.48**	.46**	.24**	.27**	.41**	.32**	.32**
Te g	.85**	.93**	.86**	1	.30**	.55**	.61**	.29**	.28**	.49**	.47**	.40**
vi	.27**	.28**	.26**	.30**	1	.41**	.27**	.11	.20*	.24**	.28**	.45**
ni	.40**	.56**	.48**	.55**	.41**	1	.43**	.20*	.14	.40**	.55**	.38**
fi	.59**	.55**	.46**	.61**	.27**	.43**	1	.17*	.17*	.29**	.59**	.30**
ga	.28**	.24**	.24**	.29**	.11	.20*	.17*	1	.38**	.35**	.14	.17*
gn	.25**	.23**	.27**	.28**	.20**	.14	.17*	.38**	1	.42**	.09	.31**
sn	.40**	.48**	.41**	.49**	.24**	.40**	.29**	.35**	.42**	1	.22**	.42**
fk	.44**	.47**	.32**	.47**	.28**	.55**	.59**	.14	.09	.22**	1	.36**
rs	.35**	.39**	.32**	.40**	.45**	.38**	.30**	.17*	.31**	.42**	.36**	1

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO gesamt; vi = verbale Intelligenz, IST 2000 R; ni = numerische Intelligenz, IST 2000 R; fi = figurale Intelligenz, IST 2000 R; ga = geteilte Aufmerksamkeit, TAP; gn = Go/Nogo, TAP; sn = Switching Numerisch, Arbeitsgedächtnis-Testbatterie; fk = Figurale Koordination, Arbeitsgedächtnis-Testbatterie; rs = Reading Span, Arbeitsgedächtnis-Testbatterie. r = bivariate Korrelation, * = $p < .05$, ** = $p < .01$.

Zur Schätzung der Konstruktreliabilität des TEMEKKO wurde versucht, ein Messmodell zu generieren, das in Anlehnung an das Aufmerksamkeitsmodell nach Moosbrugger, Goldhammer & Schweizer (2006) den Faktor Executive Attention abbilden sollte. Das Modell sollte darüber hinaus die Variablen der Arbeitsgedächtnis-Tests (Oberauer, Süß,

Schulze, Wilhelm & Wittmann, 2000) sowie der Intelligenz (I-S-T 2000R) (Amthauer, Brocke, Liepmann & Beauducel, 2001) auf latenter Ebene verorten können. Dieses Vorhaben mündete in der Modellierung verschiedener Strukturen, die entweder Schätzprobleme aufwiesen, nicht rechenbar waren oder letztlich nicht schlüssig interpretiert werden konnten.

Aus Überlegungen zu den jeweiligen Aufgabenanforderungen der verwendeten Analysevariablen, möglicher konzeptioneller Beziehungen untereinander sowie unter Einbeziehung der korrelativen Zusammenhänge wurde das nachfolgende Modell entwickelt. Inspiriert von den Befunden der Validierungsstudie 1, in der dem TEMEKKO ein hohes Maß an Executive Attention bescheinigt wurde, sollten Verfahren mit hohen Anforderungen an internaler Kontrolle innerer Prozesse als Indikatoren auf dem gleichen Faktor dienen. Die beiden TAP Subtests Geteilte Aufmerksamkeit und Go/Nogo beinhalten beide Varianzanteile, die auf selektive Aufmerksamkeit zurückgeführt werden können und Teile, die kognitive Kontrolle erfordern (Bühner, Mangels, Krumm & Ziegler, 2005). Switching Aufgaben, denen der Test Switching Numerisch (Arbeitsgedächtnis-Testbatterie) zugeordnet wird, wurden als Maß für selektive Aktivierung und Inhibition von Handlungsschemata als ein Aspekt der Überwachung und Kontrolle kognitiver Prozesse entwickelt (Oberauer, Süß, Schulze, Wittmann, 2000). Aus diesen Gründen dienten die TEMEKKO Untertests sowie die Tests Switching Numerisch, Geteilte Aufmerksamkeit und Go/Nogo als Indikatoren eines Faktors, der als exekutive Kontrolle benannt wurde.

Die beiden anderen Arbeitsgedächtnistestaufgaben Figurale Koordination sowie Reading Span gelten als Markertests des Faktors Speichern im Kontext gleichzeitiger Verarbeitung/Koordination mit jeweils unterschiedlichem Material (figural und verbal).

Beide Aufgaben erfordern die mentale Aufrechterhaltung von Information sowie die Einbettung in übergeordnete Strukturen (Relation von Punkten zu einander bzw. Wortliste in der richtigen Reihenfolge / in Relation zur Listenposition) (Oberauer, Süß, Wilhelm & Sander, 2007). Für diese Untersuchung wurden diese beiden Tests als Indikatoren für Relational Integration (ehemals Coordination) angesehen. Relational Integration bezieht sich auf die Fähigkeit, neue Beziehungen zwischen Elementen zu bilden und dabei neue strukturelle Repräsentationen zu schaffen (Waltz, Knowlton, Holyoak, Boone, Mishkin, de Menezes Santos; 1999). Die Intelligenztestvariablen wurden diesem Faktor zugedacht, weil Reasoning-Aufgaben von Probanden ebenso fordern, neue in Beziehung stehende Repräsentationen zu bilden (Oberauer et al., 2007). Eine Korrelation zwischen den latenten Variablen Relational Integration und exekutiver Kontrolle wurde aufgrund der überwiegend hohen korrelativen Zusammenhänge der Testvariablen zugelassen (und erwartet).

Darüber hinaus wurde ein Methodenfaktor aufgenommen, der für die Erfassung mittels Reaktionszeiten kontrollierte. Bei den Tests Switching Numerisch, Geteilte Aufmerksamkeit und Go/Nogo wurden Reaktionszeiten als Leistungsmaße festgehalten und luden deshalb zusätzlich auf dem Methodenfaktor. Für etwaige Einflüsse aus Materialvarianzen wurde über ausgewählte Korrelationen (Reading Span zu verbaler Intelligenz sowie figurale Intelligenz zu figuralem Subtest des TEMEKKO) kontrolliert. Mit diesen beiden Korrelationen sollten die maßgeblichen Materialeinflüsse hinreichend abgedeckt sein. Abbildung II. 1 zeigt das Messmodell zur Schätzung der Konstrukt reliabilitäten des TEMEKKO dieser Studie.

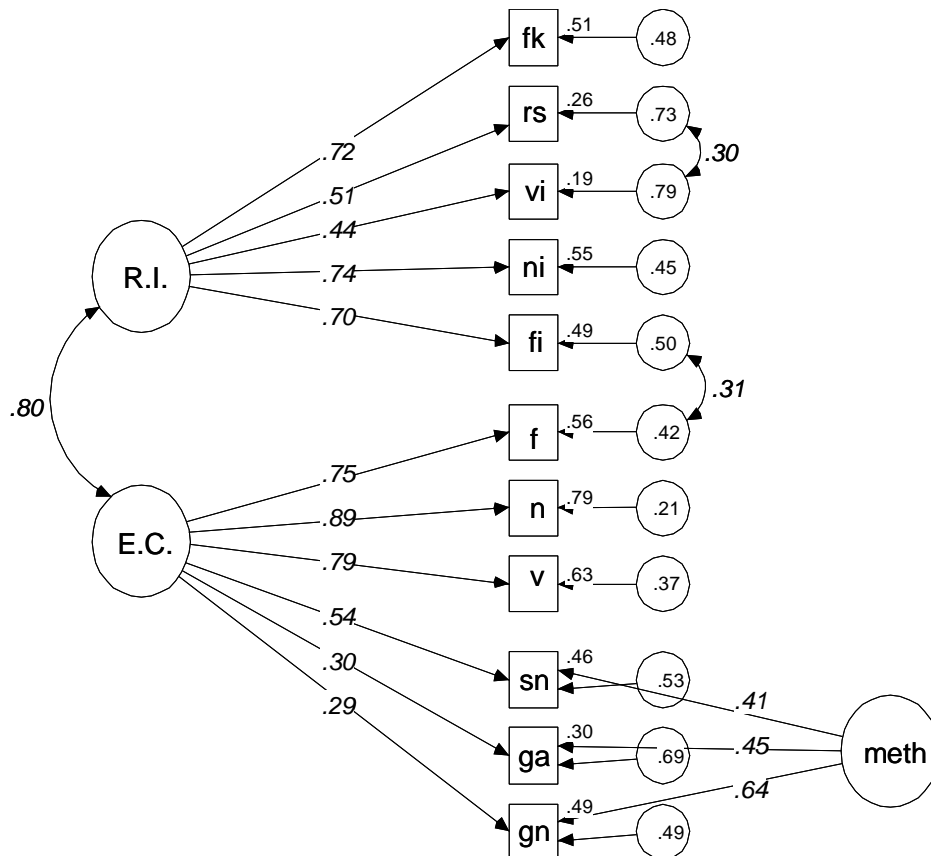


Abbildung II. 1: Messmodell zur Schätzung der Konstruktreliabilität

Anmerkungen: fk = figurale Koordination, Arbeitsgedächtnis-Testbatterie; rs = Reading Span, Arbeitsgedächtnis-Testbatterie; vi = verbale Intelligenz, IST 2000 R; ni = numerische Intelligenz, IST 2000 R; fi = figurale Intelligenz, IST 2000 R; f = TEMEKKO figural; n = TEMEKKO numerisch; v = TEMEKKO verbal; sn = Switching Numerisch, Arbeitsgedächtnis-Testbatterie; ga = Geteilte Aufmerksamkeit, TAP; gn = Go/Nogo, TAP. R.I. = Relational Integration; E.C. = Exekutive Kontrolle; meth = Methode.

Das Messmodell zeigte mit ($\chi^2[38] = 55.05$, $p = n.s.$, $CFI = .97$, $RMSEA = .06$ [.02; .09], $SRMR = .05$) einen akzeptablen Modellfit. Die Ladungen auf dem Faktor exekutive Kontrolle lagen für die Untertests des TEMEKKO bei $\lambda = .89$ (numerisch), $\lambda = .79$ (verbal) und $\lambda = .75$ (figural) für Switching Numerisch bei $\lambda = .54$ und mit $\lambda = .30$ bzw. $\lambda = .29$ für die beiden Tests zur selektiven Aufmerksamkeit (Geteilte Aufmerksamkeit bzw. Go/Nogo). Der Faktor exekutive Kontrolle wurde von den Untertests des TEMEKKO markiert. Er subsumierte Fähigkeiten, Handlungen und Gedanken in Übereinstimmung mit internalen Absichten anzuleiten, wie sie von Cohen, Botvinick und Cartner (2000) für das Konzept exekutive Kontrolle postuliert werden. Darüber hinaus steuern sie die Informationsverarbeitung und Reaktionsauswahl (response selection) in Situationen, in

denen automatisierte Mechanismen entweder nicht vorhanden, oder für die Aufgabenerfüllung nicht adäquat sind (Norman & Shallice, 1986).

Der Faktor Relational Integration wies Ladungen in Höhe von $\lambda = .74$ für numerische Intelligenz, $\lambda = .72$ für figurale Koordination, $\lambda = .70$ für figurale Intelligenz, $\lambda = .51$ für Reading Span und $\lambda = .44$ für verbale Intelligenz auf. Die Zuordnung der Verfahren zu einem Faktor schien angemessen zu sein.

Der Methodenfaktor wurde aus Reaktionszeitmaßen gebildet und erwies sich mit den substantiellen Ladungen i.H.v. $\lambda = .64$ (Go/Nogo), $\lambda = .45$ (geteilte Aufmerksamkeit) und $\lambda = .41$ (switching numerisch) als weiterer bedeutender Gruppenfaktor in der Erklärung der Kovarianzmatrix der zugrunde liegenden Analysevariablen. In diesem Sinne konnte er als speed-Faktor interpretiert werden. Alle Faktorladungen erreichten Signifikanzen auf dem 1 %o Niveau.

Die im Messmodell postulierte Korrelation zwischen den latenten Variablen exekutive Kontrolle und Relational Integration ($r = .80$) erwies sich als adäquat und war hoch signifikant ($p < .001$). Der Ordnung halber wurde geprüft, wie sich die Passung des Modells ändern würde, ließe man (1) einen perfekten Zusammenhang $r = 1$ zu bzw. (2) ginge man von einer Unabhängigkeit $r = 0$ aus. Unter beiden Annahmen wurde kein exakter Modellfit erreicht. Die beiden Faktoren teilten einen großen Anteil gemeinsamer Varianz, waren aber dennoch voneinander separierbar.

Der Einfluss von verbalem Material wurde über die Korrelation zwischen Reading Span und verbaler Intelligenz ($r = .30$) und von figuralem Material über den Zusammenhang figurale Intelligenz und TEMEKKO numerisch ($r = .31$) jeweils mit $p < .01$ signifikant.

Der Versuch, den Faktor Relational Integration nach Intelligenzaufgaben und Aufgaben der Arbeitsgedächtniskapazität zu trennen, aber auf latenter Ebene Zusammenhänge zueinander sowie zu exekutiver Kontrolle zuzulassen (Modell nicht abgebildet), führte zu Schätzproblemen und wurde verworfen. Die minderungskorrigierte Korrelation zwischen den Skalen des I-S-T 2000 R einerseits und den beiden Arbeitsgedächtnistests (figurale Koordination und Reading Span) andererseits ($r \sim 1$) wurde als weiterer Beleg für die Nähe der Konzepte verstanden (Bühner, Krumm & Pick, 2005; Bühner, Krumm, Ziegler & Plücken, 2006).

Auf Parcel-Ebene wurde die Vorhersageleistung der verwendeten Analysevariablen mit der TEMEKKO Testleistung als abhängige Variable untersucht. Mithilfe einer linearen multiplen Regressionsanalyse (Einschluss Methode) wurden die besten Prädiktoren ermittelt. Es zeigte sich, dass 46 % des erklärbaren Varianzanteils des TEMEKKO durch Intelligenz, Arbeitsgedächtnis und Aufmerksamkeit vorher gesagt werden konnten. Den größten einzigartigen Varianzanteil konnte Intelligenz erklären (11 %), gefolgt vom Parcel der Aufmerksamkeitstests (5 %). Das Parcel der Arbeitsgedächtniskapazitätstests war in der Zusammensetzung durch figurale Koordination und Reading Span von untergeordneter Bedeutung zur Vorhersage des TEMEKKO über Intelligenz und Aufmerksamkeit hinaus, obwohl der gemeinsame Varianzanteil mit dem Kriterium ($r^2_{simple} = .27$) auf dem 1 % Niveau signifikant war. Dieser gemeinsame Anteil enthielt offenkundig auch Anteile anderer Prädiktoren, die für deren einzigartige Vorhersage am TEMEKKO stärker maßgeblich waren. In dieser Prädiktion blieb der Einfluss der Methodenvarianz (speed) unberücksichtigt. Tabelle II. 9 zeigt die Ergebnisse der Analyse.

Tabelle II. 9: Regressionsanalyse mit TEMEKKO als Kriterium auf Parcel-Ebene

	β	r^2_{simple}	$r^2_{semi-partial}$
INT	.46 ***	.40 **	.11 ***
AUF	.25 ***	.21 **	.05 ***
AG	.12	.27 **	.05

AV = TEMEKKO, $R^2 = .46$

Anmerkungen: INT = Parcel Intelligenz = mean (verbale, numerische, figurale Intelligenz, I-S-T 2000 R); AG = Parcel Arbeitsgedächtnis = mean (figurale Koordination, Reading Span, Arbeitsgedächtnis-Testbatterie); AUF = Parcel Aufmerksamkeit = mean (switching numerisch, Arbeitsgedächtnis Testbatterie; geteilte Aufmerksamkeit, Go/Nogo, TAP). β = standardisierter Regressionskoeffizient, r^2_{simple} = quadrierte bivariate Korrelation, $r^2_{semi-partial}$ = quadrierte Semipartialkorrelation, AV = abhängige Variable; R^2 = korrigierter Determinationskoeffizient; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Die Aufgabenbearbeitung des TEMEKKO erfordert per se kein schlussfolgerndes Denken, wie es beispielsweise für den I-S-T 2000 R postuliert wird. Daher wurden aufgrund der funktionalen Enge zwischen Konzentration und Aufmerksamkeit die Indikatoren des Parcels Aufmerksamkeit einer weiteren linearen multiplen Regressionsanalyse (Einschluss Methode) mit der TEMEKKO Testleistung als Kriterium unterzogen.

Hier erwies sich der Test Switching Numerisch ($r^2_{semi-partial} = .16$) als einzig maßgeblicher Prädiktor in der Vorhersage des erklärbaren Varianzanteils der TEMEKKO Testleistung (24 %). Switching numerisch erfordert kognitive Flexibilität (wechseln zwischen mentalen Sets) und gilt als Indikator für Supervision/Shifting und Bearbeitungsgeschwindigkeit. Die beiden TAP Untertests Geteilte Aufmerksamkeit und Go/Nogo erfordern beide kognitive Kontrolle sowie selektive Aufmerksamkeit, was möglicherweise die gemeinsamen Varianzanteile am Kriterium (jeweils 8 %) erklärte. Für die Prädiktion des TEMEKKO eigneten sie sich aber aufgrund ihrer geringen spezifischen Vorhersagekraft (2 % bzw. < 1 %) nicht. Tabelle II. 10 zeigt die Befunde.

Tabelle II. 10: Regressionsanalyse mit TEMEKKO als Kriterium

	β	r^2_{simple}	$r^2_{semi-partial}$
sn	.42 ***	.24 ***	.16 ***
ga	.12	.08 ***	.02
gn	.06	.08 **	< .01

AV = TEMEKKO, $R^2 = .24$

Anmerkungen: sn = Switching Numerisch, Arbeitsgedächtnis-Testbatterie; ga = Geteilte Aufmerksamkeit, TAP; gn = Go/Nogo, TAP; β = standardisierter Regressionskoeffizient, r^2_{simple} = quadrierte bivariate Korrelation, $r^2_{semi-partial}$ = quadrierte Semipartialkorrelation, AV = abhängige Variable; R^2 = korrigierter Determinationskoeffizient; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Zusammenfassung der Validitätsschätzungen

Das aus konzeptionellen Überlegungen entwickelte Messmodell bestand aus zwei korrelierenden, aber separierbaren Faktoren auf Funktionsebene. Der Faktor Exekutive Kontrolle wurde in erster Linie vom numerischen Untertests des TEMEKKO ($\lambda = .89$) markiert, gefolgt vom TEMEKKO verbal ($\lambda = .79$) und TEMEKKO figural ($\lambda = .75$). Der zweite Faktor Relational Integration wurde am höchsten durch die Indikatoren numerische Intelligenz ($\lambda = .74$) und figurale Koordination ($\lambda = .72$) geprägt. Die hohe Korrelation der beiden Faktoren ($r = .80$) verdeutlichte den engen funktionalen Zusammenhang der Konzepte; nichtsdestotrotz waren sie von einander abzugrenzen. Der Faktor Relational Integration konnte in dieser Untersuchung nicht weiter in seine konzeptionellen Bereiche (Intelligenz und Arbeitsgedächtniskapazität) separiert werden.

Für die Vorhersage des TEMEKKO auf Parcel-Ebene zeigte sich Intelligenz mit 11 % einzigartigem Varianzanteil als bester Prädiktor des erklärbaren Varianzanteils des Kriteriums (46 %). 5 % der erklärbaren Kriteriumsvarianz konnte auf die Maße der Aufmerksamkeitstests zurückgeführt werden. Von diesen wiederum erwies sich die

Shifting-Anforderung (Switching Numerisch) als einziger Prädiktor mit 16 % einzigartigem Anteil an der erklärbaren TEMEKKO-Varianz (24 %) als maßgeblich.

Der signifikante Zusammenhang zwischen den figuralen Untertests des TEMEKKO und der figuralen Intelligenz ($r = .31$) machte einen Einfluss der figuralen Inhaltsdomäne deutlich, wenngleich die Höhe des Zusammenhangs als eher gering eingestuft wurde. Bedeutsame verbale und/oder numerische Materialeinflüsse zeigten sich bei den TEMEKKO Untertests nicht. Der TEMEKKO wurde demnach als relativ material-unabhängig angesehen.

Validierungsstudie 3

Der Datensatz für die Validierungsstudie 3 stammte aus einer Untersuchung eines Diplomanden der Philipps-Universität Marburg, Fachbereich Psychologie.

1 Testdurchführung

Die Erhebung der Daten fand im Zeitraum von Juni bis November 2002 in zwei Gruppenräumen des Psychologischen Instituts der Phillips-Universität Marburg statt. Die insgesamt 29 Testungen wurden grundsätzlich in Gruppen durchgeführt (13 in Zweier- und Dreiergruppen, 15 in Gruppen zu 5 bis 10 Personen sowie eine Einzeltestung). Die Instruktionen wurden vom Testleiter vorgetragen und lagen den Probanden in schriftlicher Form vor. Die Generalinstruktion beinhaltete Informationen über das Ziel, den Umfang und den Aufbau der Untersuchung sowie über den generellen Umgang mit den Testmaterialien. Nachdem die Versuchspersonen auf einem Erfassungsblatt Angaben zu Testpersonenummer, Geschlecht, Alter, Studien(haupt)fach bzw. Beruf sowie Datum der Testung und Angaben über übliche Schlafzeiten und –dauer machten, schloss der Einleitungsteil und die eigentliche Durchführung der Tests wurde gestartet.

Die Reihenfolge der Testdarbietung wurde bewusst gewählt. Die Testung erfolgte in drei Blöcken zu jeweils ca. einer $\frac{3}{4}$ Stunde mit Pausen von 15 Minuten Länge dazwischen, um Ermüdungseffekte auf die Konzentrationsleistung möglichst gering zu halten. Vor dem Hintergrund, dass die Bearbeitung des Berliner-Intelligenz-Struktur-Tests (BIS-4-S) die Konzentrationsleistung beeinflussen könnte, wurden die Verfahren in den dritten Testblock

eingepplant. Des weiteren wurde etwaigen Transfereffekten von numerischen auf figurale Konzentrationstests (Bühner, 2001; Westhoff & Dewald, 1990) Rechnung getragen, indem die Konzentrationstests mit figuralem Material in den ersten beiden Blöcken jeweils vor den zahlengebundenen Verfahren angeordnet wurden. Die Durchführungszeiten der Verfahren INKA und TEMEKKO waren mit mehr als 20 Minuten geplant und wurden daher in unterschiedlichen Blöcken und jeweils an den Schluss der Testblöcke 1 und 2 gesetzt. Der Test MWT-B eignete sich aufgrund seiner einfachen Instruktion besonders gut als Einstiegstest und verhinderte durch seine unbegrenzte Bearbeitungszeit eine gewisse Unruhe während der Testung. Die übrigen Tests wurden nach zeitökonomischen Gesichtspunkten den drei Testblöcken zugeordnet.

2 Stichprobe

Der Datensatz umfasste 128 Personen im Alter von 19 bis 30 Jahren ($M = 21.8$ Jahre; $SD = 2.39$ Jahre). Das mittlere Alter der 97 weiblichen Versuchspersonen (75.8 %) betrug 21.4 Jahre ($SD = 2.14$ Jahre), das der 31 männlichen Probanden 23.0 Jahre ($SD = 2.73$ Jahre).

126 der 128 Probanden aus dieser Stichprobe waren Studenten der Universität. Die übrigen 2 Personen waren Erzieherinnen. Nähere Angaben über Zugehörigkeit der Studierenden zu ihren Studienfächern lagen nicht vor.

Die Anwerbung der Versuchspersonen erfolgte direkt über Aushänge in den Aufenthaltsräumen der Philipps-Universität Marburg. Als Entlohnung für die Teilnahme an der Testung erhielten die Studenten ($N = 79$) 3.5 Versuchspersonenstunden. Die übrigen

Testpersonen ($N = 49$) bekamen eine telefonische Rückmeldung über ihre Konzentrations- und Intelligenzleistungen.

3 Versuchsmaterial

Nach Angaben der jeweiligen Testautoren wurden die Verfahren TEMEKKO, d2, KLT-R, INKA, FAIR und Rev.-T als Maße der Konzentration angesehen. Die Tests MWT-B und BIS galten als Intelligenztests. Der Test GFT wurde dem Konzept Figurenwahrnehmung zugeschrieben, der Motoriktest den motorischen Fähigkeiten und der Test DRE3 stand für einfache Rechenleistungen. Aufgrund der Untersuchung von Schmidt-Atzert, Bühner, Enders (2006) wurde die Zuordnung der Verfahren zu einzelnen Konstrukten abgeändert. Demnach wurden die Subtests X-Größer (XG), Zahlen-Symbol-Test (ZS), Buchstaben Durchstreichen (BD), Klassifizieren Wörter (KW), Teil Ganzes (TG) aus dem BIS, sowie der GFT und Motoriktest dem Konstrukt Konzentration zugeschrieben. INKA galt als besserer Indikator für Intelligenz als für Konzentration. Die Tests KLT, DRE sowie die BIS Subtests Unvollständige Wörter (UW), Zahlenpaare (ZP) und Sinnvoller Text (ST) wurden als Repräsentanten für das Konstrukt Gedächtnis verwendet. In Tabelle III.1 ist der genaue zeitliche Ablauf der Testung dargestellt.

Tabelle III.1: Ablaufplan der Testung

Testverfahren	Untertest	Dauer	Art der Bearbeitung
MWT-B		5 min.	Paper-Pencil
Test d2		7 min 40 sec.	Paper-Pencil

Tabelle III.1 wird fortgesetzt

KLT-R (Kurzform)		13 min	Paper-Pencil
INKA		21 min	Paper-Pencil
<i>Pause</i>		15 min	
GFT Gleiche Figuren Test		2 min 15 sec	Paper-Pencil
FAIR		10 min	Paper-Pencil
Rev.-T. Kurzform		6 min	Paper-Pencil
BIS-4	KW Klassifizieren Wörter	1 min	Paper-Pencil
BIS-4	ZS Zahlen-Symbol-Test	1 min 30 sec	Paper-Pencil
TEMEKKO		21 min 30 sec	Paper-Pencil
<i>Pause</i>		15 min	
Motoriktest		2 min 10 sec	Paper-Pencil
DRE 3		4 min 30 sec	Paper-Pencil
BIS-4	UW Unvollständige Wörter	1 min 20 sec	Paper-Pencil
BIS-4	OG Orientierungsgedächtnis	3 min 10 sec	Paper-Pencil
BIS-4	ZN Zahlenreihen	5 min	Paper-Pencil
BIS-4	EF Eigenschaften/Fähigkeiten	3 min 15 sec	Paper-Pencil
BIS-4	AN Analogien	3 min	Paper-Pencil
BIS-4	XG X-Größer	2 min	Paper-Pencil
BIS-4	WA Wortanalogien	2 min 45 sec	Paper-Pencil
BIS-4	LO Layout	4 min	Paper-Pencil
BIS-4	ZP Zahlenpaare	4 min 30 sec	Paper-Pencil
BIS-4	TM Tatsache/Meinung	2 min	Paper-Pencil
BIS-4	BD Buchstaben durchstreichen	1 min 10 sec	Paper-Pencil
BIS-4	SC Schätzen	4 min 15 sec	Paper-Pencil
BIS-4	ST Sinnvoller Text	3 min	Paper-Pencil
BIS-4	DR Divergentes Rechnen	2 min 30 sec	Paper-Pencil
BIS-4	CH Chargo	4 min 20 sec	Paper-Pencil
BIS-4	TG Teil Ganzes	2 min	Paper-Pencil

3.1 TEMEKKO

Die detaillierte Beschreibung des Verfahrens findet sich in Abschnitt 3 (Entwicklung des TEMEKKO) im Manual.

3.2 Mehrfach-Wahl-Wortschatz-Intelligenz-Test (MWT)

(Lehrl, 1995). Die Versionen des MWT Verfahrens sind Leistungstests zur Erfassung des allgemeinen Intelligenzniveaus, insbesondere des Niveaus der kristallisierten Intelligenz. In der vorliegenden Untersuchung wurde die Testform B eingesetzt. Als Operationalisierung von Wortschatzkenntnissen stellt dieser Test insgesamt 37 Aufgaben an die Versuchspersonen. Die Aufgaben sind nach Schwierigkeit ansteigend, zeilenweise angeordnet sind. In jeder Aufgabe werden vier fiktiv konstruierte, in der deutschen Sprache nicht existente Wörter sowie ein existierendes deutschsprachiges Wort dargeboten. Dieses eine existierende Wort ist zu markieren. Die Bearbeitung des Tests ist zeitlich nicht limitiert, die Bearbeitung der 37 Items beläuft sich üblicherweise auf ca. 5 Minuten. Als Leistungsmaß wurde die Anzahl der richtig beantworteten Aufgaben gewertet.

3.3 Aufmerksamkeits-Belastungs-Test d2

(Brickenkamp, 2002). Für eine detaillierte Beschreibung des Verfahrens sei auf Abschnitt 3.2 (Aufmerksamkeits-Belastungs-Test d2) im Anhang I (Validierungsstudie 1) verwiesen.

3.4 Konzentrations-Leistungs-Test (KLT-R)

(Lukesch & Mayrhofer, 2001). Der KLT-R erfasst die Quantität als auch die Qualität der Dauerbeanspruchung und des Leistungsverlaufs einer Testperson. Der Test misst, im Unterschied zu Tests der Kurzzeitanspannung (z.B. Test d2), die Langzeitanspannung. Der KLT-R existiert in zwei Schwierigkeitsstufen. Die leichtere Version (KLT-R 4-6) ist für Schüler der 4. bis 6. Schulstufe geeignet. Die hier eingesetzte schwierigere Fassung (KLT-R 6-13) ist in der 6. bis 13. Schulstufe und bei Erwachsenen anwendbar. Für beide Schwierigkeitsstufen wurden zwei echte Parallelversionen entwickelt. Die in der vorliegenden Studie verwendete Version ist die Form A. Das Verfahren besteht aus neun Aufgabenblöcken, die spaltenweise angeordnet sind. Jede Spalte enthält 20 Rechenaufgaben. Eine Rechenaufgabe besteht aus der Addition und Subtraktion von zwei mal drei einstelligen Zahlen. Die beiden Zwischenergebnisse müssen memoriert werden. Ist das erste Zwischenergebnis größer als das zweite, so muss das zweite vom ersten subtrahiert werden. Im umgekehrten Fall müssen die beiden Zwischenergebnisse addiert werden. Die Endergebnisse sind in Kästchen einzutragen. Um die Bearbeitungszeit der gesamten Testung zu verringern, wurden den Probanden lediglich die ersten vier Blöcke (von ansonsten neun Blöcken) dargeboten. Als Leistungsmaß wurde die Anzahl der richtig beantworteten Aufgaben erfasst.

3.5 Inventar Komplexer Aufmerksamkeit (INKA)

(Heyde, 2000). Das Verfahren erfasst die Konzentrationsleistung im Sinne komplexer Aufmerksamkeitsleistungen über die Güte und Intensität der Informationsverarbeitung. Der Aspekt Ausdauer spielt eine untergeordnete Rolle, vielmehr sollen Aufgaben mit

unterschiedlichem Schwierigkeitsgrad unter hoher Konzentration gelöst werden. Der Testbogen besteht aus insgesamt 18 Aufgaben, deren Schwierigkeiten im Laufe der Bearbeitung kontinuierlich ansteigen. Die Aufgabenreize sind Reihen zufällig kombinierter Konsonanten. Auf der linken Seite neben den Buchstabenreihen befinden sich einzelne Konsonanten oder Konsonantenpaare, die mit Hilfe einer separaten Umwandlungstabelle von den Probanden in andere Konsonanten umkodiert werden müssen. Die umkodierten Konsonanten/-paare sind in den Buchstabenreihen zu entdecken. Der links neben dem entdeckten Konsonanten/-paar stehende Buchstabe ist als Target rechts neben der Suchreihe zu notieren. Eine Aufgabe ist dann richtig gelöst, wenn alle Targets in der richtigen Reihenfolge eingetragen sind. Als Leistungsmaß wurde die Anzahl der richtig beantworteten Aufgaben erfasst.

3.6 Gleiche-Figuren-Test (GFT)

(unveröffentlichtes Verfahren, Bühner & Schmidt-Atzert). Dieser Test existiert bislang lediglich in einer unveröffentlichten Fassung und misst die Figurenwahrnehmung der Testpersonen. Aufgrund der begrenzten Bearbeitungszeit von 15 Sekunden kann auf die Schnelligkeit geschlossen werden, mit der zwei gleiche Objekte unter Ähnlichen wahrgenommen werden. Es ist davon auszugehen, dass die Wahrnehmung von Objekten von deren Form, Textur und Größe abhängig ist (Goldstein, 1989). Daher wurden die drei Subtests des GFT anhand dieser Merkmale konzipiert. Das Reizmaterial jedes Untertests besteht aus fünf Zeilen mit je vier Aufgaben (Kästchen). In der Mitte jedes Kästchens befindet sich ein Objekt, das von sechs Objekten kreisförmig umschlossen ist. Eines der sechs Objekte gleicht dem mittig Dargestellten. „Gleich“ bedeutet im Subtest 1 die gleiche Form, im Subtest 2 die gleiche Textur und in Subtest 3 die gleiche Größe. Der Proband

muss jedes Kästchen durchsehen und die beiden jeweils gleichen Objekte durch einen Strich miteinander verbinden. Die Anzahl der richtig beantworteten Aufgaben wurde als Leistungsmaß angesehen.

3.7 Frankfurter Aufmerksamkeitsinventar (FAIR)

(Moosbrugger & Oehlschlägel, 1996). Das Verfahren zur Aufmerksamkeits- und Konzentrationsdiagnostik misst gerichtete Aufmerksamkeit als Fähigkeit zur konzentrierten, d.h. genauen und schnellen Diskrimination visuell ähnlicher Zeichen unter gleichzeitiger Ausblendung aufgabenirrelevanter Information. Das Instrument ist ein Test zur Aufmerksamkeits- und Konzentrationsdiagnostik in den Dimensionen Leistung, Qualität und Kontinuität. Gerichtete Aufmerksamkeit als Fähigkeit wird in zwei Testbögen erfasst. Jeder Bogen besteht aus 16 Zeilen zu je 20 Items. Ein Item setzt sich entweder aus einem Kreis mit zwei oder drei Punkten, oder aus einem Quadrat mit zwei oder drei Punkten zusammen. Innerhalb des Kreises bzw. des Quadrats variiert die Anordnung der Punkte, was für die Beantwortung der Aufgabe aber irrelevant ist. Als Targets sind Kreise mit drei Punkten und Quadrate mit zwei Punkten definiert. Die Probanden sollen so schnell und fehlerfrei als möglich eine durchgezogene Linie von links nach rechts unter den Distraktoren ziehen. Die Targets sind durch einen Zacken nach oben (zum Item) zu markieren. Die Linie soll dabei nicht unterbrochen werden. Die Bearbeitungszeit ist mit drei Minuten für jeden Testbogen begrenzt. Das Leistungsmaß, das in die Berechnungen eingeflossen ist, war die Anzahl der richtig beantworteten Items beider Testbögen.

3.8 Revisions Test (Rev. T.)

(Marschner, 1980). Der allgemeine Leistungstest zur Untersuchung anhaltender Konzentration bei geistiger Tempoarbeit besteht aus 15 Testzeilen zu je 44 Items. Für die vorliegende Untersuchung wurde eine kürzere Fassung verwendet, die in 15 Testzeilen pro Zeile lediglich 22 Aufgaben beinhaltet. Ein Item besteht aus drei einstelligen Zahlen, die untereinander stehen. Die Probanden müssen prüfen, ob die Summe der beiden oberen Zahlen gleich der unteren Zahl ist (oder nicht). Ist die Summe gleich der unteren Zahl, so ist das Item mit einem Haken zu versehen. Sind die Werte unterschiedlich, ist das Item mit einem senkrechten Strich durchzustreichen. Die Bearbeitungszeit pro Zeile wird mit 15 Sekunden vorgegeben. Die Anzahl der richtig bearbeiteten Items galt als Leistungsmaß.

3.9 Berliner Intelligenz-Struktur-Test (BIS-4 -S)

(Jäger, Süß, Beauducel, 1997). Der BIS (Form 4) stellt eine umfangreiche Testbatterie dar, mit der Intelligenzleistungen besonders breit und vielfältig erfasst werden. Die für diese Untersuchung verwendete Kurzform (S) des Verfahrens unterscheidet sich zur Grundform durch eine verkürzte Generalinstruktion. Das zugrunde liegende Modell wurde auf sehr aufwändige Weise entwickelt. Ca. 2000 in der Literatur auffindbare Intelligenzaufgabentypen wurden gesammelt und daraus 191 Itemblöcke zu je 98 Aufgaben extrahiert. 500 Probanden bearbeiteten diese repräsentative Stichprobe aller Intelligenzaufgaben. Die so erhobenen Daten wurden analysiert und daraus ein hierarchisches, bimodales Intelligenzstrukturmodell entwickelt. Die beiden Modalitäten umfassen zum einen die Art des Aufgabenmaterials (verbal, numerisch, figural) und zum anderen die Operationen (kognitiven Prozesse), die zur Aufgabenbearbeitung erforderlich sind (Bearbeitungs-

geschwindigkeit, Gedächtnis, Einfallsreichtum und Verarbeitungskapazität). Der BIS existiert in einer Originalform (BIS-4) und einer Kurzfassung (BIS-4-S). Die Kurzfassung besteht aus Aufgaben mit den höchsten Trennschärfen unter Beibehaltung der BIS-spezifischen Bandbreite und hat eine im Vergleich zur Langversion verkürzte Generalinstruktion. Die Subtests (16) des BIS lassen sich mit Ausnahme des Tests Unvollständige Wörter (UW) zu einem allgemeinen Intelligenzquotienten verrechnen. Darüber hinaus lassen sich die einzelnen Subtests anhand ihrer jeweiligen Zuordnung zu den operationsgebundenen bzw. zu den materialgebundenen Facetten im Strukturmodell verorten. Alle für die vorliegende Untersuchung verwendeten BIS-4 Subtests wurden mittels Schablonen ausgewertet.

3.9.1 Klassifizieren Wörter (KW)

Zuordnung zu den bimodalen Facetten: Bearbeitungsgeschwindigkeit verbal

Eine dreispaltige Liste besteht aus 75 Substantiven, von denen 40 eine Pflanze bezeichnen. Innerhalb von 30 Sekunden muss die Versuchsperson so viele Pflanzenwörter wie möglich durchstreichen. Die Anordnung der Targets ist zufällig. Die Anzahl der richtig durchgestrichenen Worte wurde als Leistungsmaß erfasst.

3.9.2 Zahlen-Symbol-Test (ZS)

Zuordnung zu den bimodalen Facetten: Bearbeitungsgeschwindigkeit figural

In jeder der vier Testzeilen stehen 17 Items. Jedes Item besteht aus zwei Quadraten, die untereinander angeordnet sind. Das obere Quadrat enthält eine Ziffer von 1 bis 9. In einer

separat zum Testbogen vorgelegten Umwandlungstabelle ist jeder Ziffer von 1 bis 9 ein Symbol zugeordnet. Der Proband soll in das untere Quadrat jeweils das entsprechende Symbol (ermittelt aus der Umwandlungstabelle) eintragen. Die Items werden ohne Auslassung von links nach rechts zeilenweise durchgearbeitet. Als Leistungsmaß galt die Anzahl der richtig bearbeiteten Items.

3.9.3 Unvollständige Wörter (UW)

Zuordnung zu den bimodalen Facetten: Bearbeitungsgeschwindigkeit verbal

In drei Spalten zu je 19 Items befinden sich 57 deutschsprachige Substantive, in denen jeweils ein Buchstabe fehlt. Dieser Buchstabe ist vom Probanden auf dem dafür vorgesehenen Auslassungsstrich zu ergänzen. Die Items werden spaltenweise von oben nach unten bearbeitet. Die Anzahl der richtig beantworteten Items entsprach dem Leistungsmaß.

3.9.4 Orientierungsgedächtnis (OG)

Zuordnung zu den bimodalen Facetten: Merkfähigkeit figural

Zunächst wird den Versuchsteilnehmern ein Ausschnitt aus einem Stadtplan im DIN-A4 Format 90 Sekunden lang präsentiert. Die Instruktion verlangt, sich die schwarz gekennzeichneten Gebäude und Grundstücke einzuprägen. Nach einer Merkphase sollen die Probanden aus dem Gedächtnis auf einem gleichen Stadtplanausschnitt ohne schwarze Markierungen die zuvor gekennzeichneten Gebäude und Grundstücke mit einem Kreuz

einzeichnen. Die Bearbeitungszeit ist für diese Aufgaben auf 100 Sekunden beschränkt. Leistungsmaß war die Anzahl der richtig gesetzten Kreuze.

3.9.5 Zahlenreihen (ZN)

Zuordnung zu den bimodalen Facetten: Verarbeitungskapazität numerisch

In neun Zeilen (Items) sind jeweils fünf bis sieben Zahlen angeordnet, deren Abfolge jeweils eine rechnerische Regelmäßigkeit zugrunde liegt. Die jeweils nächste der Regelmäßigkeit entsprechende Zahl ist zu ergänzen. Die Anzahl der richtig fortgesetzten Zahlenreihen galt als Leistungsmaß.

3.9.6 Eigenschaften/Fähigkeiten (EF)

Zuordnung zu den bimodalen Facetten: Einfallsreichtum verbal

In der Übungsaufgabe (ca. 1 Minute) sollen Probanden so viel wie möglich Eigenschaften und Fähigkeiten listen, die ein guter Richter *nicht* haben sollte. Für die eigentliche Testaufgabe sollen so viel wie möglich Eigenschaften und Fähigkeiten aufnotiert werden, die ein guter Verkäufer *nicht* haben sollte. Als Leistungsmaß wurden die Anzahl der Nennungen gewertet.

3.9.7 Analogien (AN)

Zuordnung zu den bimodalen Facetten: Verarbeitungskapazität figural

In acht Zeilen (Items) befinden sich im linken Teil jeweils zwei Figuren, die zueinander in einer Beziehung stehen sowie eine einzelne Figur und ein Fragezeichen. Aus den fünf im

rechten Teil der Zeile dargebotenen Figuren soll jene ausgesucht werden, die anstelle des Fragezeichens eingesetzt werden muss, damit die gleiche Relation zwischen der einzelnen und der statt dem Fragezeichen ergänzten Figur entsteht, wie zwischen den ersten beiden Figuren. Die richtige Figur muss aus den fünf dargebotenen angestrichen werden. Als Leistungsmaß wurde die Anzahl der richtig bearbeiteten Items gewertet.

3.9.8 X-Größer (XG)

Zuordnung zu den bimodalen Facetten: Bearbeitungsgeschwindigkeit numerisch

Bei den Übungsaufgaben zu diesem Subtest müssen die Probanden alle Zahlen, die um sieben größer sind als die vorangegangene Zahl, durchstreichen. Bei den Testaufgaben sind jene Zahlen durchzustreichen, die um drei größer sind als die Vorangegangenen. Der Testbogen enthält 130 ein- und zweistellige Zahlen in zehn Zeilen zu je 13 Zahlen. Die Bearbeitungsrichtung entspricht der üblichen Schreibrichtung, von links nach rechts nacheinander. Als Leistungsmaß galt die Anzahl der richtig markierten Zahlen.

3.9.9 Wortanalogien (WA)

Zuordnung zu den bimodalen Facetten: Verarbeitungskapazität verbal

Der Aufbau und der Bearbeitungsmodus der Aufgaben entsprechen jenen des Subtests Analogien (AN), mit dem Unterschied, dass Substantive anstelle von Figuren dargeboten werden. Die Anzahl der richtig bearbeiteten Items diente als Leistungsmaß.

3.9.10 Layout (LO)

Zuordnung zu den bimodalen Facetten: Merkfähigkeit numerisch

Die Übungsphase dieser Aufgabe sieht vor, dass Probanden Logos für einen Laden namens *Brillenwerner* skizzenhaft entwerfen sollen. Die Zeichnungen sollen Hinweise auf die Ware und den Ladenbesitzer beinhalten. Als Beispiele sind bereits zwei der neun Übungskästchen mit Grafiken ausgefüllt. Für die eigentliche Testaufgabe stehen 12 leere Kästchen zur Verfügung, in die möglichst viele Logos für den Laden *Jupp's Fahrradcenter* zu entwerfen sind. Die Anzahl der Skizzen wurden als Leistungsmaß gewertet.

3.9.11 Zahlenpaare (ZP)

Zuordnung zu den bimodalen Facetten: Merkfähigkeit numerisch

Innerhalb von zwei Minuten sollen die Probanden 12 untereinander stehende Zahlenpaare, die auf einer DIN-A4 Seite präsentiert werden, memorieren. Auf der nächsten Testseite werden den Versuchspersonen nur die links präsentierte Zahl des jeweiligen Zahlenpaares präsentiert sowie fünf weitere Zahlen, von denen eine die ursprüngliche zweite (rechte) Zahl des Paares ist. Die Reihenfolge der ursprünglichen Zahlenpaare ist für die Suche nach der richtigen rechten Zahl verändert. Die Bearbeitungszeit für die Wiedererkennung der Zahlen ist auf zwei Minuten begrenzt. Als Leistungsmaß galt die Anzahl der richtig bearbeiteten Items.

3.9.12 Tatsache/Meinung (TM)

Zuordnung zu den bimodalen Facetten: Verarbeitungskapazität verbal

Auf dem Testbogen stehen untereinander 16 Aussagesätze, neben denen sich jeweils die Großbuchstaben *T* und *M* befinden. Diese Sätze sollen dahingehend beurteilt werden, ob deren Inhalt einer Tatsache oder einer Meinung entspricht. Je nachdem kreuzt der Proband *T* oder *M* an. Die Anzahl der richtig angekreuzten Aufgaben diene als Leistungsmaß.

3.9.13 Buchstaben Durchstreichen (BD)

Zuordnung zu den bimodalen Facetten: Bearbeitungsgeschwindigkeit figural

13 Buchstabenreihen bilden die Aufgabe, bei der so schnell als möglich alle *x* durchgestrichen werden müssen. Die Zeilen werden in Schreibrichtung von links nach rechts nacheinander bearbeitet. Als Leistungsmaß wurde die Anzahl der richtig durchgestrichenen Targets gewertet.

3.9.14 Schätzen (SC)

Zuordnung zu den bimodalen Facetten: Verarbeitungskapazität numerisch

Das Testmaterial besteht aus sieben Mehrfach-Wahlaufgaben, die untereinander angeordnet sind. Dargeboten werden Rechenaufgaben, die sich aus zwei oder drei mehrstelligen Zahlen und den Rechenzeichen +, - und x zusammensetzen. Aufgrund der Größe der Zahlenwerte ist es üblicherweise nicht möglich, die Aufgaben durch schnelles Kopfrechnen zu lösen. Gleichwohl besteht die Möglichkeit, eine vereinfachende Regel

herzuleiten, mit der eine der fünf angebotenen Lösungsmöglichkeiten richtigerweise identifiziert werden kann. Die Anzahl der richtig beantworteten Items wurde als Leistungsmaß gewertet.

3.9.15 Sinnvoller Text (ST)

Zuordnung zu den bimodalen Facetten: Merkfähigkeit verbal

Den Probanden wird ein aus 85 Wörtern bestehender Text vorgegeben, aus dem sie sich so viele Angaben wie möglich einprägen sollen. Inhaltlich befasst sich der Text damit, welche Unternehmen Direktor Salzmann zu seinem Ruhestand gratulieren, wo die Abschiedsfeier stattfindet und wer daran teilnimmt. Nach einer Einprägephase von einer Minute werden den Versuchspersonen 22 Detailfragen zum Text gestellt, die sie schriftlich innerhalb von zwei Minuten beantworten sollen. Als Leistungsmaß ging die Anzahl der richtig reproduzierten Wörter in die Berechnungen ein.

3.9.16 Divergentes Rechnen (DR)

Zuordnung zu den bimodalen Facetten: Einfallsreichtum numerisch

Bei dieser Aufgabe sollen die Probanden möglichst viele verschiedene Zahlenkombinationen aufschreiben, die bei aufeinanderfolgender Multiplikation und Addition die Summe 60 ergeben (Beispiel: $4 \times 10 + 20 = 60$). Werden mehr als fünf Kombinationen durch die selbe Lösungsstrategie erbracht, so werden nur fünf Kombinationen bei der Auswertung angerechnet (Beispiel $0 \times 1 + 60$; $0 \times 2 + 60$ usw.). Generell galt die Zahl der Lösungsstrategien als wertbares Leistungsmaß.

3.9.17 Chargo (CH)

Zuordnung zu den bimodalen Facetten: Verarbeitungskapazität figural

Das Reizmaterial besteht aus sechs Reihen von Zeichnungen, die folgerichtig um zwei weitere Zeichnungen ergänzt werden sollen. Vier Zeichnungen werden für jedes Item als Auswahlösungen angeboten. Als Leistungsmaß wurde die Anzahl der richtig bearbeiteten Items erfasst.

3.9.18 Teil-Ganzes (TG)

Zuordnung zu den bimodalen Facetten: Einfallsreichtum numerisch

Auf dem Testbogen werden den Probanden drei Spalten mit jeweils 19 untereinander stehenden Substantiven dargeboten. Alle Worte, unter denen ein Wort steht, das aus materieller Sicht einen Teil des darüberstehenden Wortes darstellt (z.B. Teil: Schornstein, Ganzes: Fabrik), sind anzukreuzen. Die Spalten werden von oben nach unten bearbeitet. Die Anzahl der richtig markierten Worte wurde als Leistungsmaß gewertet.

3.10 Motoriktest

(Schmidt-Atzert, Bühner, 1997). Dieses Verfahren wurde aus der Auswertungsschablone des Tests d2 (Brickenkamp, 1987) entwickelt und erfordert vom Probanden motorische Schnelligkeit. In den 14 Zeilen der Schablone stehen jeweils 47 Kästchen, von denen 359 schwarz ausgefüllt sind und die übrigen 299 einen weißen Kreis enthalten. Diese Kreise sollen so schnell als möglich durchgestrichen werden. Die Bearbeitung des Tests erfolgt in Schreibrichtung von links nach rechts. Die Bearbeitungszeit für jede Zeile wurde

mit fünf Sekunden vorgegeben. Als Leistungsmaß galt die Anzahl der richtig bearbeiteten Items.

3.11 Diagnostischer Rechentest für 3. Klassen (DRE3)

(Samstag, Sander, Schmidt & Ingenkamp, 1971). Insgesamt besteht das Verfahren aus 44 Items. Da die letzten 4 Items Textaufgaben sind und neben den Fertigkeiten zu addieren, subtrahieren, multiplizieren und dividieren auch verbale textgebundene Anforderungen an die Probanden stellen, wurden diese vier Items bei der Testung vernachlässigt. Die übrigen 40 Aufgaben bestehen aus zwei Spalten mit je 20 Items. Der Proband muss diese spaltenweise von oben nach unten bearbeiten. Die Ergebnisse der Berechnungen sind auf einem Testbogen einzutragen. Dieser Leistungstest ermöglicht anhand von Normen eine objektive Feststellung und Beurteilung der Rechenfertigkeit und des Rechenverständnisses. Als Leistungsmaß galt die Anzahl der richtig berechneten Items.

4 Statistische Analysen

4.1 Deskriptiva

Nachfolgende Tabelle III.2 zeigt einen Überblick über die Rohdaten der Analysevariablen.

Tabelle III.2: Deskriptive Statistiken der Testvariablen

Tests	<i>N</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>	<i>r_{age}</i>	<i>r_{gender}</i>
Te fig	128	24.0	73.0	46.4	8.5	0.33	0.23	-.11	.00
Te num	128	12.0	78.0	46.5	12.0	0.23	0.57	.12	.12
Te verb	128	16.0	66.0	43.7	8.9	-0.29	0.80	-.04	-.13
Te g	128	76.0	212.0	136.5	5.4	0.26	0.23	.00	.01

Tabelle III.2 wird fortgesetzt

KW	128	14.0	43.0	29.3	4.8	-0.36	0.80	.08	-.19*
ZS	126	26.0	68.0	42.0	7.5	0.16	0.21	-.12	-.35**
UW	128	19.0	55.0	35.5	6.5	0.03	0.35	-.03	-.08
OG	126	4.0	24.0	13.4	4.4	0.00	-0.77	.03	-.00
ZN	128	0.0	9.0	5.2	2.2	-0.11	-0.80	-.04	.11
EF	128	0.0	30.0	11.6	4.7	0.79	2.30	.03	-.12
AN	128	0.0	8.0	3.8	1.6	0.23	-0.03	-.03	-.03
XG	128	10.0	44.0	25.3	6.8	0.27	-0.11	.23**	.24**
WA	128	0.0	8.0	4.2	2.0	-0.05	-0.55	-.00	.08
LO	128	1.0	9.0	4.5	1.6	0.83	0.62	-.09	-.06
ZP	128	1.0	12.0	6.9	2.8	0.17	-0.70	-.03	-.25**
TM	128	1.0	16.0	10.9	3.2	-0.40	-0.44	.02	.01
BD	128	41.0	87.0	55.6	8.3	1.00	1.62	-.05	-.07
SC	128	0.0	7.0	4.2	1.8	-0.39	-0.46	-.07	.02
ST	128	3.0	19.0	11.3	3.3	-0.17	-0.41	-.07	-.19*
CH	128	0.0	6.0	2.9	1.6	0.00	-0.79	-.05	.11
TG	127	5.0	21.0	13.3	2.7	-0.08	0.54	-.04	-.12
DR	128	4.0	25.0	11.3	4.3	0.51	0.21	.00	.08
INKA	127	2.0	18.0	10.7	3.3	-0.45	-0.25	-.05	.08
MWT	128	15.0	37.0	29.0	3.7	-0.54	0.81	.22*	.19*
d2	126	105.0	299.0	196.5	37.6	0.21	0.11	-.02	-.10
FAIR	128	115.0	319.0	228.4	38.9	0.02	0.04	.01	-.15
KLT	127	9.0	65.0	29.1	12.0	0.72	-0.10	-.08	.04
GFT F	128	0.0	20.0	11.3	2.6	-0.06	2.90	-.23**	-.03
GFT M	128	9.0	20.0	14.7	2.4	0.13	0.21	-.15	-.10
GFT G	128	2.0	16.0	8.5	2.1	0.18	0.66	-.19*	-.05
Rev	128	122.0	309.0	212.3	33.0	0.26	0.52	.07	-.08
Motor	127	97.0	227.0	169.3	26.0	-0.03	0.12	.10	-.10
Rechen	127	1.0	34.0	18.0	5.9	0.16	0.11	-.03	.16

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO; KW = Klassifizieren Wörter, BIS-4; ZS = Zahlen-Symbol-Test, BIS-4; UW = Unvollständige Wörter, BIS-4; OG = Orientierungsgedächtnis, BIS-4; ZN = Zahlenreihen, BIS-4; EF = Eigenschaften/Fähigkeiten, BIS-4; AN = Analogien, BIS-4; XG = X-Größer, BIS-4; WA = Wortanalogien, BIS-4; LO = Layout, BIS-4; ZP = Zahlenpaare, BIS-4; TM = Tatsache/Meinung, BIS-4; BD = Buchstaben durchstreichen, BIS-4; SC = Schätzen, BIS-4; ST = Sinnvoller Text, BIS-4; CH = Chargo, BIS-4; TG = Teil Ganzes, BIS-4; DR = Divergentes Rechnen, BIS-4; INKA = Inventar Komplexer Aufmerksamkeit; MWT = Mehrfach-Wortschatz-Intelligenz-Test; d2 = Test d2; FAIR = Frankfurter Aufmerksamkeitsinventar; KLT = Konzentrations-Leistungs-Test; GFT F = Gleiche Figuren Test, Form; GFT M = Gleiche Figuren Test, Muster; GFT G = Gleiche Figuren Test, Größe; Rev = Revisionstest; Motor = Motorik-Test; Rechen = Rechentest. r = bivariate Korrelation, * = $p < .05$, ** = $p < .01$.

Die Verteilungsformen der einzelnen Testvariablen wurden anhand von Histogrammen und zugehörigen Normalverteilungskurven sowie ihrer numerischen Kennwerte für Schiefe und Kurtosis in Augenschein genommen und zeigten keine Auffälligkeiten.

Streudiagramme, die über die einzelnen Analysevariablen kreuzweise angefertigt wurden, zeigten ebenso keine Besonderheiten, sodass keine Fälle als Ausreißer eliminiert wurden.

4.2 Zusammenhang mit Alter und Geschlecht

Die Korrelationen mit den z-transformierten Variablen und Alter und Geschlecht zeigten generell eher schwache Zusammenhänge, bis auf wenige Ausnahmen. Eine Umpolung einzelner Analysevariablen war nicht erforderlich, weil keine Reaktionszeiten enthalten waren. Nachfolgende Tabelle zeigt die korrelativen Zusammenhänge der Analysevariablen vor und nach Auspartialisierung der Variablen Alter und Geschlecht, diese waren auf dem 1 %o Niveau signifikant. Die optische Prüfung der Zusammenhänge mittels Streudiagrammen zeigte durchgängig einen positiven linearen Trend.

Tabelle III. 3: Zusammenhang z-transformierte Testvariablen und standardisierte Residuen

z \ Res	Te fig	Te num	Te verb	Te g	KW	ZS	UW	OG	ZN	EF	AN	XG	WA
Te fig	.993												
Te num		.989											
Te verb			.992										
Te g				1.00									
KW					.970								
ZS						.938							
UW							.997						
OG								1.00					
ZN									.991				
EF										.990			
AN											.999		
XG												.956	
WA													.997

Anmerkungen.: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO; KW = Klassifizieren Wörter, BIS-4; ZS = Zahlen-Symbol-Test, BIS-4; UW = Unvollständige Wörter, BIS-4; OG = Orientierungsgedächtnis, BIS-4; ZN = Zahlenreihen, BIS-4; EF = Eigenschaften/Fähigkeiten, BIS-4; AN = Analogien, BIS-4; XG = X-Größer, BIS-4; WA = Wortanalogien, BIS-4.

z \ Res	LO	ZP	TM	BD	SC	ST	CH	TG	DR	INKA	MWT	d2
LO	.995											
ZP		.967										
TM			1.00									
BD				.997								
SC					.997							
ST						.981						
CH							.991					
TG								.993				
DR									.997			
INKA										.994		
MWT											.967	
d2												.995

Anmerkungen: LO = Layout, BIS-4; ZP = Zahlenpaare, BIS-4; TM = Tatsache/Meinung, BIS-4; BD = Buchstaben durchstreichen, BIS-4; SC = Schätzen, BIS-4; ST = Sinnvoller Text, BIS-4; CH = Chargo, BIS-4; CH = Chargo, BIS-4; TG = Teil Ganzes, BIS-4; DR = Divergentes Rechnen, BIS-4; INKA = Inventar Komplexer Aufmerksamkeit; MWT = Mehrfach-Wortschatz-Intelligenz-Test; d2 = Test d2.

z \ Res	FAIR	KLT	GFT F	GFT M	GFT G	Rev	Motor	Rechen
FAIR	.988							
KLT		.995						
GFT F			.972					
GFT M				.988				
GFT G					.983			
Rev						.992		
Motor							.986	
Rechen								.983

Anmerkungen: FAIR = Frankfurter Aufmerksamkeitsinventar; KLT = Konzentrations-Leistungs-Test; GFT F = Gleiche Figuren Test, Form; GFT M = Gleiche Figuren Test, Muster; GFT G = Gleiche Figuren Test, Größe; Rev = Revisionstest; Motor = Motorik-Test; Rechen = Rechentest.

4.3 Schätzung fehlender Werte

Zur Schätzung der fehlenden Werte anhand des EM-Algorithmus ergab der MCAR-Test nach Little mit $\chi^2[248] = 266.23$, n.s., dass die MCAR-Nullhypothese angenommen werden konnte. Die korrelativen Zusammenhänge sowie die grafische Überprüfung der standardisierten Residuen und jenen um fehlende Werte Ergänzten waren erwartungsgemäß zufriedenstellend. Es lagen lediglich 2 fehlende Werte bei den Variablen

d2 und den beiden Subtests ZS und OG aus dem BIS vor, sowie jeweils nur 1 Wert in den Variablen KLT, INKA, Motorik, Rechnen und dem BIS-Subtest TG vor, i.e. max. 1.6 % bei $N = 128$.

4.4 Prüfung auf Gleichheit der Testformen und Materialarten

Zur Überprüfung der Strukturgleichungsmodelle (vgl. Abbildung 3 in Abschnitt 4.1.2 (Prozedere der Datenbearbeitung) des Manuals) seien zunächst die Ergebnisse des Mardia-Tests zur Prüfung der Annahme der multivariaten Normalverteilung berichtet. (Tempowerte: multivariate Kurtosis = 9.08, c.r. = 5.24, $p = < .001$ bzw. Fehler-Prozentwerte: multivariate Kurtosis = 38.63, c.r. = 22.30, $p = < .001$). Aufgrund der Verletzung der Annahme wurde die Prüfung der Strukturgleichungsmodelle unter Berücksichtigung der Bollen-Stine-Bootstrap-Korrektur vorgenommen. Hinsichtlich der von West, Finch und Curran (1995) postulierten Grenzen für Schiefe (< 2) und Kurtosis (< 7) war festzuhalten, dass die Variablen der Tempowerte innerhalb dieser Grenzen lagen. Die Fehler-Prozent-Variablen hingegen überschritten diese Grenzen wie folgt: Die figural-Fehler-B-Prozente lagen mit 2.09 für die Schiefe knapp über und mit 6.30 für die Kurtosis knapp innerhalb des Limits. Verbal-Fehler-A-Prozente überschritten die Schiefe mit 2.21 und die Kurtosis mit 7.35.

Die Prüfung der Messmodelle auf Passung der Daten zeigte für Tempowerte (Modell a) ($\chi^2 [16] = 28.10$, $p = \text{n.s.}$, $CFI = .98$, $RMSEA = .08$ [.02; .12], $SRMR = .05$) als auch für die Fehler-Prozent-Werte (Modell b) ($\chi^2 [16] = 10.95$, $p = \text{n.s.}$, $CFI = 1.00$, $RMSEA = .00$ [.00; .05], $SRMR = .04$) akzeptable Modellfits auf. Die beiden Testformen A und B, sowie die

drei Materialarten (figural, numerisch und verbal) konnten als essentiell-parallele Tests angesehen werden und erfassten dementsprechend das selbe Konstrukt gleich gut.

Die Untersuchung der Materialarten (figural, numerisch und verbal) hinsichtlich ihrer Unterschiede zueinander anhand von Kovarianzanalysen mit Alter und Geschlecht als Kovariate ergaben zunächst für den 3-fach Vergleich sowie für den Vergleich figural : numerisch statistisch signifikante Unterschiede. Die Nullhypothese (Gleichheit der Materialien) konnte in diesen Fällen nicht beibehalten werden. Die Untersuchung kritischer hypothetischer Mittelwertsdifferenzen unter den Bedingungen der Prüfung auf Gleichheit und den strengen Anpassungskriterien der Teststärke nach Bonferroni-Holm ergab drei statistisch und praktisch bedeutsame Vergleiche. Naturgemäß blieben die beiden vorhin benannten Vergleiche signifikant und zusätzlich erlangte der paarweise Vergleich figural : verbal Signifikanz. Die Effektgrößen des 3-fach Vergleichs ($\eta^2 = .03$) sowie der Paarung figural : verbal ($\eta^2 = .02$) sind als gering einzustufen, jene des Vergleichs figural : numerisch ($\eta^2 = .05$) im mittleren Bereich. Diese Befunde veranlassten die Bereitstellung separater Normen für die jeweiligen Materialarten. Die Ergebnisse der Berechnungen zeigt Tabelle III.4.

Tabelle III.4: Prüfung der Materialien auf Gleichheit

Material	F_{emp}	p	η^2	power'	F_{krit}
figural : numerisch	6.261	.014	.048	.983	0.353
numerisch : verbal	1.335	.250	.011	.950	1.529
figural : verbal	2.567	.112	.020	.975	0.273
figural : numerisch : verbal	3.534	.031	.027	.950	1.639

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert

Die Prüfung der Formen A und B auf deren Unterschiedlichkeit bzw. Gleichheit anhand von Kovarianzanalysen nach analogen Berechnungen ergab weder vor noch nach dem

Prozedere des Abgleichs anhand des kleinstmöglich angenommenen Effektes statistische oder praktische Bedeutung. Die Testformen A und B konnten in allen Fällen als gleich angesehen werden. Die Ergebnisse hierzu sind in Tabelle III.5 gelistet.

Tabelle III.5: Prüfung der Testformen auf Gleichheit

Testformen	F_{emp}	p	η^2	power'	F_{krit}
A : B (figural)	1.159	.284	.009	.983	2.295
A : B (numerisch)	0.089	.766	.001	.975	6.782
A : B (verbal)	0.015	.902	<.001	.950	4.884

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert

4.5 Reliabilitäten

Die Reliabilitätsschätzungen der Tempowerte der in dieser Untersuchung verwendeten Verfahren sind in Tabelle III.6 dargestellt.

Tabelle III.6: Reliabilitätsmaße der verwendeten Verfahren

Test	N	r_{tt}
TEMEKKO figural	128	.89 ^a
TEMEKKO numerisch	128	.94 ^a
TEMEKKO verbal	128	.91 ^a
TEMEKKO gesamt	128	.96 ^a
MWT-B ^d	128	.75 ^b
Test d2 ^d	128	.94 ^c
KLT-R ^d	128	.91 ^b
INKA ^d	128	.79 ^b
GFT-Form ^d	128	.94 ^b
GFT-Muster ^d	128	.93 ^b
GFT-Größe ^d	128	.84 ^b
FAIR ^d	128	.86 ^c
Rev.-Test ^d	128	.97 ^c
Motorik-Test ^d	128	.96 ^c
DRE 3 ^d	128	.92 ^c

Tabelle III.6 wird fortgesetzt

BIS-4 ^d , KW Klassifizieren Wörter	128	.99 ^b
BIS-4 ^d , ZS Zahlen-Symbol-Test	128	.98 ^b
BIS-4-S ^d , UW Unvollständige Wörter	128	.97 ^b
BIS-4-S ^d , OG Orientierungsgedächtnis	128	.61 ^b
BIS-4-S ^d , ZN Zahlenreihen	128	.78 ^c
BIS-4-S ^d , AN Analogien	128	.48 ^b
BIS-4-S ^d , XG X-Größer	128	.92 ^b
BIS-4-S ^d , WA Wortanalogien	128	.66 ^b
BIS-4-S ^d , ZP Zahlenpaare	128	.73 ^b
BIS-4-S ^d , TM Tatsache/Meinung	128	.79 ^b
BIS-4-S ^d , BD Buchstaben durchstreichen	128	.78 ^b
BIS-4-S ^d , SC Schätzen	128	.60 ^c
BIS-4-S ^d , ST Sinnvoller Text	128	.73 ^b
BIS-4-S ^d , CH Chargo	128	.60 ^b
BIS-4-S ^d , TG Teil Ganzes	128	.71 ^b

^a Parallelestreliabilität

^b Split-Half-Reliabilität

^c Cronbach α

^d aus Schmidt-Atzert, Bühner & Enders (2006)

4.6 Validität

Tabelle III.7 zeigt zunächst die Korrelationsmatrix der Testvariablen dieser Validierungsstichprobe.

Auffällig hohe Zusammenhänge der TEMEKKO Variablen ließen sich hieraus mit dem Frankfurter Aufmerksamkeitsinventar (FAIR), dem Revisions Test sowie mit dem Test d2 erkennen. Diese Verfahren gelten als Tests zur Erfassung von Konzentration und Aufmerksamkeit. Von den Aufgaben des Intelligenztests BIS-4 zeigten mehrere Untertests hohe Zusammenhänge mit dem TEMEKKO, i.e. ZP, ZS, AN oder ZN, insbesondere XG.

Tabelle III.7: Korrelationsmatrix der Analysevariablen

<i>r</i>	Te fig	Te num	Te verb	Te g	KW	ZS	UW	OG	ZN	EF	AN
Te fig	1	.65**	.57**	.83**	.32**	.39**	.17	.27**	.29**	.17	.36**
Te num	.65**	1	.69**	.92**	.25**	.40**	.29**	.18*	.39**	.17	.39**
Te verb	.57**	.69**	1	.86**	.40**	.38**	.34**	.22*	.40**	.05	.39**
Te g	.83**	.92**	.86**	1	.36**	.45**	.31**	.25**	.42**	.15	.43**
KW	.32**	.25**	.40**	.36**	1	.18*	.47**	.15	.19*	.15	.19*
ZS	.39**	.40**	.38**	.45**	.18*	1	.29**	.16	.34**	.04	.25**
UW	.17	.29**	.34**	.31**	.47**	.29**	1	.04	.28**	.01	.12
OG	.27**	.18*	.22*	.25**	.15	.16	.04	1	.14	.03	.06
ZN	.29**	.39**	.41**	.42**	.19*	.34**	.28**	.14	1	-.01	.42**
EF	.17	.17	.05	.15	.15	.04	.01	.03	-.01	1	-.09
AN	.36**	.39**	.39**	.43**	.19*	.25**	.12	.06	.42**	-.10	1
XG	.38**	.55**	.50**	.58**	.39**	.36**	.50**	.22*	.40**	.11	.25**
WA	.16	.25**	.41**	.31**	.36**	.00	.21*	.19*	.25**	.08	.34**
LO	.15	.10	.15	.15	.03	.20*	.04	.18*	.16	.37**	.08
ZP	.40**	.40**	.40**	.46**	.19*	.22*	.26**	.31**	.26**	-.01	.26**
TM	.27**	.36**	.37**	.39**	.47**	.17	.17	.13	.25**	.14	.28**
BD	.24**	.06	.16	.16	.20*	.31**	.19*	.16	.05	-.04	.08
SC	.25**	.30**	.24**	.31**	.21*	.17	.11	.13	.36**	-.00	.17
ST	.03	.15	.18*	.15	.36**	.13	.32**	.06	.15	.10	.14
CH	.28**	.28**	.25**	.31**	.05	.17	.08	.19*	.31**	-.12	.37**
TG	.24**	.19*	.28**	.26**	.43**	.12	.27**	.03	.16	.12	.15
DR	.33**	.41**	.28**	.40**	.12	.28**	.26**	.10	.33**	.22*	.18*
INKA	.25**	.34**	.44**	.39**	.22*	.22*	.26**	.14	.31**	-.06	.32**
MWT	.23**	.18*	.29**	.26**	.53**	-.02	.41**	.11	.21*	.15	.14
d2	.34**	.49**	.47**	.50**	.27**	.50**	.31**	.23**	.36**	.08	.31**
FAIR	.51**	.46**	.58**	.58**	.29**	.43**	.27**	.34**	.24**	.02	.32**
KLT	.29**	.45**	.43**	.46**	.37**	.35**	.50**	.19*	.47**	.02	.33**
GFT	.42**	.39**	.43**	.47**	.27**	.44**	.39**	.11	.07	-.05	.27**
Rev.-T.	.43**	.54**	.56**	.59**	.37**	.62**	.45**	.19*	.35**	.05	.23**
Motorikt.	.36**	.24**	.24**	.32**	.01	.25**	.12	.07	.13	.15	.10
Rechent.	.28**	.42**	.42**	.43**	.36**	.31**	.42**	.14	.35**	.13	.24**

Tabelle III.7 wird fortgesetzt

<i>r</i>	XG	WA	LO	ZP	TM	BD	SC	ST	CH	TG	DR
Te fig	.38**	.16	.15	.40**	.27**	.24**	.25**	.03	.28**	.24**	.33**
Te num	.55**	.25**	.10	.40**	.36**	.06	.30**	.15	.28**	.19*	.41**
Te verb	.50**	.41**	.15	.40**	.37**	.16	.24**	.18*	.25**	.28**	.28**
Te g	.58**	.31**	.15	.46**	.39**	.16	.31**	.15	.31**	.26**	.40**
KW	.39**	.36**	.03	.19*	.47**	.20*	.21*	.36**	.05	.43**	.12
ZS	.36**	.00	.20*	.22*	.17	.31**	.17	.13	.17	.12	.28**
UW	.50**	.21*	.04	.26**	.17	.19*	.11	.32**	.08	.27**	.26**
OG	.22*	.19*	.18*	.31**	.13	.16	.13	.06	.19*	.03	.10
ZN	.40**	.25**	.16	.26**	.25**	.05	.36**	.15	.31**	.16	.33**
EF	.11	.08	.37**	-.01	.14	-.04	-.00	.10	-.12	.12	.22*
AN	.25**	.34**	.08	.26**	.28**	.08	.17	.14	.37**	.15	.18*
XG	1	.11	.08	.30**	.30**	.24**	.33**	.25**	.16	.30**	.50**
WA	.11	1	.23**	.20*	.43**	-.00	.08	.16	.36**	.28**	.02
LO	.08	.23**	1	.06	.21*	.11	.07	.05	.12	-.05	.19*
ZP	.30**	.20*	.06	1	.19*	.05	.20*	.25**	.30**	.23*	.21*
TM	.30**	.43**	.21*	.19*	1	-.06	.33**	.11	.24**	.25**	.17
BD	.24**	-.00	.11	.05	-.06	1	.19*	-.09	.17	.12	.13
SC	.33**	.08	.07	.20*	.33**	.19*	1	-.01	.28**	.24**	.21*
ST	.25**	.16	.05	.25**	.11	-.09	-.01	1	.04	.15	.12
CH	.16	.36**	.12	.30**	.24**	.17	.28**	.04	1	.17	.00
TG	.30**	.28**	-.05	.23*	.25**	.12	.24**	.15	.17	1	.08
DR	.50**	.02	.19*	.21*	.17	.13	.21*	.12	.00	.08	1
INKA	.23*	.27**	.13	.25**	.28**	.20*	.27**	.05	.28**	.10	.16
MWT	.30**	.30**	.15	.14	.38**	.01	.12	.20*	.13	.22*	.19*
d2	.49**	.18*	.26**	.18*	.25**	.30**	.22*	.11	.26**	.08	.32**
FAIR	.40**	.15	.11	.33**	.23**	.37**	.19*	.10	.21*	.22*	.32**
KLT	.65**	.23**	.02	.38**	.27**	.12	.31**	.32**	.16	.27**	.54**
GFT	.33**	.12	.07	.32**	.06	.41**	.10	.12	.20*	.12	.14
Rev.-T.	.76**	.08	.16	.30**	.28**	.35**	.28**	.19*	.11	.25**	.45**
Motorikt.	.25**	-.03	.02	.12	-.04	.21*	.02	-.04	-.07	.10	.25**
Rechent.	.58**	.18*	.19*	.35**	.40**	.15	.39**	.29**	.20*	.21*	.50**

Tabelle III.7 wird fortgesetzt

<i>r</i>	INKA	MWT	d2	FAIR	KLT	GFT	Rev.-T.	Motorikt.	Rechent.
Te fig	.25**	.23**	.34**	.51**	.29**	.42**	.43**	.36**	.28**
Te num	.34**	.18*	.49**	.46**	.45**	.39**	.54**	.24**	.42**
Te verb	.44**	.29**	.47**	.58**	.43**	.43**	.56**	.24**	.42**
Te g	.39**	.26**	.50**	.58**	.46**	.47**	.59**	.32**	.43**
KW	.22*	.53**	.27**	.29**	.37**	.27**	.37**	.01	.36**
ZS	.22*	-.02	.50**	.43**	.35**	.44**	.62**	.25**	.31**
UW	.26**	.41**	.31**	.27**	.50**	.39**	.45**	.12	.42**
OG	.14	.11	.23**	.34**	.19*	.11	.19*	.07	.14
ZN	.31**	.21*	.36**	.24**	.47**	.07	.35**	.13	.35**
EF	-.06	.15	.08	.02	.02	-.05	.05	.15	.13
AN	.32**	.14	.31**	.32**	.33**	.27**	.23**	.09	.24**
XG	.23*	.30**	.49**	.40**	.65**	.33**	.76**	.25**	.58**
WA	.27**	.30**	.18*	.15	.23**	.12	.08	-.03	.18*
LO	.13	.15	.26**	.11	.02	.07	.16	.02	.19*
ZP	.25**	.14	.18*	.33**	.38**	.32**	.30**	.12	.35**
TM	.28**	.38**	.25**	.23**	.27**	.06	.30**	-.04	.40**
BD	.20*	.01	.30**	.37**	.12	.41**	.35**	.21*	.15
SC	.27**	.12	.23*	.19*	.31**	.10	.28**	.02	.39**
ST	.05	.20*	.11	.10	.32**	.12	.19*	-.04	.29**
CH	.28**	.13	.26**	.21*	.16	.20*	.11	-.07	.20*
TG	.10	.22*	.08	.22*	.27**	.12	.25**	.10	.21*
DR	.16	.19*	.32**	.32**	.54**	.14	.45**	.25**	.50**
INKA	1	.21*	.39**	.35**	.34**	.22*	.27**	.04	.34**
MWT	.21*	1	.18*	.15	.24**	.15	.22*	-.02	.31**
d2	.39**	.18*	1	.49**	.33**	.38**	.56**	.24**	.41**
FAIR	.35**	.15	.49**	1	.36**	.45**	.50**	.30**	.36**
KLT	.34**	.24**	.33**	.36**	1	.23**	.62**	.11	.65**
GFT	.22*	.15	.38**	.45**	.23**	1	.46**	.20*	.22*
Rev.-T.	.27**	.22*	.56**	.50**	.62**	.46**	1	.33**	.55**
Motorikt.	.04	-.02	.24**	.30**	.11	.20*	.33**	1	.08
Rechent.	.34**	.31**	.41**	.36**	.65**	.22*	.55**	.08	1

Anmerkungen.: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch, Te verb = TEMEKKO verbal; Te g = TEMEKKO gesamt; KW = Klassifizieren Wörter; ZS = Zahlen-Symbol-Test; UW = Unvollständige Wörter; OG = Orientierungsgedächtnis; ZN = Zahlenreihen; EF = Eigenschaften/Fähigkeiten; AN = Analogien; XG = X-Größer; WA = Wortanalogien; LO = Layout; ZP = Zahlenpaar; TM = Tatsache/Meinung; BD = Buchstaben durchstreichen; SC = Schätzen; ST = Sinnvoller Text; CH = Chargo; TG = Teil Ganzes; DR = Divergentes Rechnen; INKA = Inventar Komplexer Aufmerksamkeit; MWT = Mehrfach-Wortschatz-Intelligenz-Test; d2 = Test d2; FAIR = Frankfurter Aufmerksamkeitsinventar; KLT = Konzentrations-Leistungs-Test; GFT = Gleiche Figuren Test, Parcel (Form, Muster, Größe); Rev.-T. = Revisionstest; Motorikt. = Motorik-Test; Rechent. = Rechentest. *r* = bivariate Korrelation; * = $p < .05$, ** = $p < .01$.

Zur Schätzung der Konstruktvalidität in dieser Studie wurde ein Messmodell generiert, das die Zusammenhänge der TEMEKKO Testleistung mit Intelligenz und Konzentration, abbildet. Darüber hinaus wurde für die Inhaltsfacetten (verbal, numerisch und figural) kontrolliert. Die Auswahl der Testvariablen, die in die Konstrukte Intelligenz und Konzentration als Indikatoren einfließen, erfolgte gemäß den Befunden von Schmidt-Atzert, Bühner und Enders (2006). Sie untersuchten die Struktur von Konzentrationstestleistungen im Zusammenhang mit anderen Konstrukten wie Intelligenz und Gedächtnis. Als Indikatoren der latenten Variablen Konzentration und Intelligenz wurden jene Testverfahren aus der Untersuchung von Schmidt-Atzert et al. (2006) verwendet, die den jeweiligen Faktor und die zugehörige Inhaltsdomäne markierten. Für Konzentration wurden demnach der Revisions-Test (Rev. Test) für die numerische Facette, das Frankfurter Aufmerksamkeitsinventar (FAIR) als figuraler und der Subtest Klassifizieren Wörter (KW) aus dem Berliner Intelligenz-Strukturtests (BIS-4-S) als verbaler Indikator verwendet. In die latente Variable Intelligenz flossen die Untertests Zahlenreihen (ZN) (numerisch), Analogien (AN) (figural) sowie Tatsache/Meinung (TM) (verbal) des BIS-4-S ein. Aufgrund theoretischer Überlegungen zu den Konstrukten sowie der Korrelationsmatrix wurden (hohe) Zusammenhänge auf latenter Ebene erwartet, sodass im Messmodell (Abbildung III.1) Korrelationen zugelassen wurden.

Dieses Modell zeigte mit $\chi^2 [15] = 20.79$, $p = \text{n.s.}$, $CFI = .99$, $RMSEA = .06$ [.00; .11], $SRMR = .03$ einen akzeptablen Modellfit. Zunächst war zu erkennen, dass die Zusammenhänge der TEMEKKO Testleistungen in den Inhaltsdomänen niedrig und nicht signifikant ausfielen. So lagen die Korrelationen des numerischen Teils des TEMEKKO bei $r = -.01$ und $r = .09$ mit dem jeweils domänenanalogen Test bei Intelligenz und Konzentration, für den figuralen Untertest des TEMEKKO bei $r = .07$ mit Intelligenz (figural) und $r = .13$ mit Konzentration (figural) sowie bei $r = .07$ und $r = .17$ für die

verbale TEMEKKO Testleistung mit den entsprechenden verbalen Indikatoren für Intelligenz und Konzentration. Anhand dieser niedrigen und nicht signifikanten Werte für den TEMEKKO kann von einer (relativen) Materialunabhängigkeit gesprochen werden.

Die Korrelationen der Konstrukte auf latenter Ebene wiesen einen engeren Zusammenhang (minderungskorrigiert) des TEMEKKO mit Konzentration ($r = .86$) als mit Intelligenz ($r = .77$) auf, wobei beide Werte auf dem 1 %o Niveau signifikant waren, wenn für die Materialarten (verbal, numerisch und figural) kontrolliert wurde. Ohne Berücksichtigung der Materialfaktoren (Modell nicht abgebildet), stieg der Zusammenhang zwischen dem TEMEKKO und Konzentration auf $r = .90$, wohingegen die Korrelation zu Intelligenz nahezu gleich blieb ($r = .78$).

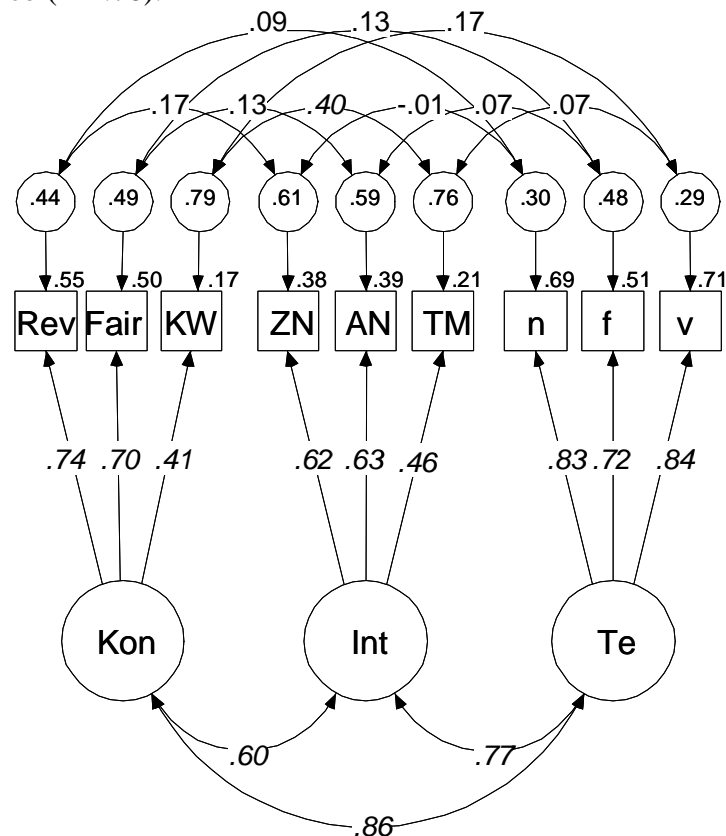


Abbildung III.1: Modell zur Prüfung der Konstruktvalidität des TEMEKKO

Anmerkungen: Rev = Revisionstest, Fair = Frankfurter Aufmerksamkeitsinventar, KW = BIS-4, Klassifizieren Wörter; ZN = BIS-4, Zahlenreihen; AN = BIS-4, Analogien; TM = BIS-4, Tatsache/Meinung; n = TEMEKKO numerisch, f = TEMEKKO figural; v = TEMEKKO verbal. Kon = Konzentration; Int = Intelligenz; Te = TEMEKKO.

Der Befund aus der Prüfung des Strukturgleichungsmodells wurde nachfolgend aus einem anderen Blickwinkel aufgegriffen. Hierfür wurden die Parcels Konzentration und Intelligenz aus zahlreicheren Testvariablen - als im o.g. Modell - gebildet und in einer linearen multiplen Regressionsanalyse als Prädiktoren zur Vorhersage des TEMEKKO verwendet. Als zusätzlicher Prädiktor wurde ein Aggregat aus Gedächtnistests in die Analyse aufgenommen. Die Zusammensetzung der Aggregate erfolgte wiederum aufgrund vorangegangener Analysen von Schmidt-Atzert, Bühner und Enders (2006). Hierfür wurden die jeweils stärksten (i.S.v. signifikanten) Indikatoren zu Parcels verrechnet. Konzentration wurde durch Mittelung folgender Tests dargestellt: Revisions Test (Rev. Test), Test d2, Frankfurter Aufmerksamkeitsinventar (FAIR), Gleiche Figuren Test (GFT), Motoriktest sowie die Subtests X-Größer (XG), Zahlen-Symbol-Test (ZS), Buchstaben Durchstreichen (BD), Klassifizieren Wörter (KW) und Teil Ganzes (TG) aus dem Berliner Intelligenz-Strukturtest (BIS-4-S). Die Tests Inventar Komplexer Aufmerksamkeit (INKA) sowie die Subtests des BIS-4-S Zahlenreihen (ZN), Schätzen (SC), Tatsache/Meinung (TM), Wortanalogien (WA), Analogien (AN) und Chargo (CH) bildeten das Aggregat der Intelligenz. Zusätzlich wurde ein weiteres Parcel aus Gedächtnisleistungstests als Prädiktor gebildet. Dieses setzte sich aus den Verfahren Konzentrations-Leistungs-Test (KLT), Diagnostisches Rechnen (DRE) und die BIS-4-S Untertests Unvollständige Wörter (UW), Zahlenpaare (ZP) und Sinnvoller Text (ST) zusammen. Bei dieser Regressionsanalyse konnte nicht um den Einfluss von Materialvarianzanteilen kontrolliert werden.

Anhand der breiteren Zusammensetzung der Parcels - breiter als die Aggregate im Strukturgleichungsmodell - konnte Konzentration als stärkster Prädiktor 15 % des erklärbaren Varianzanteils (54 %) des TEMEKKO vorhersagen. Intelligenz zeigte mit 7 % einen ebenso hoch signifikanten einzigartigen Anteil an der Vorhersage, wenngleich auch von geringerer Ausprägung. Die Gedächtnisfunktion hingegen schien keine maßgebliche

Fähigkeit zu sein, die zur Erklärung der TEMEKKO Testleistung entscheidend beitragen konnte ($r^2_{\text{semi-partial}} < .001$), obwohl der gemeinsame Varianzanteil (26 %) mit dem Kriterium hoch signifikant war. Dieser beinhaltete offenkundig auch Anteile anderer Prädiktoren, die in deren einzigartigen Vorhersagekraft Ausdruck fanden. Nachfolgende Tabelle III.8 zeigt die Ergebnisse der Regressionsanalyse.

Tabelle III.8: Regressionsanalyse mit TEMEKKO als Kriterium auf Parcel-Ebene

	β	r^2_{simple}	$r^2_{\text{semi-partial}}$
Konz	.51 ***	.47 ***	.15 ***
Int	.32 **	.33 ***	.07 ***
Mem	.04	.26 ***	<.001

AV = TEMEKKO, R^2 .54

Anmerkungen: Konz = Parcel aus (Rev., BIS XG, d2, FAIR, BIS ZS, BIS BD, GFT Parcel, Motorik, BIS KW, BIS TG), Int = Parcel aus (INKA, BIS ZN, BIS SC, BIS TM, BIS WA, BIS AN, BIS CH), Mem = Parcel aus (KLT, DRE, BIS UW, BIS ZP, BIS ST). β = standardisierter Regressionskoeffizient, r^2_{simple} = quadrierte bivariate Korrelation, $r^2_{\text{semi-partial}}$ = quadrierte Semipartialkorrelation, AV = abhängige Variable; R^2 = korrigierter Determinationskoeffizient; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Um die Vorhersagen an Intelligenz näher zu beleuchten, wurde eine weitere multiple Regressionsanalyse mit Intelligenz als Kriterium und dem TEMEKKO, Konzentration und Gedächtnis auf Parcel-Ebene als Prädiktoren durchgeführt (Ergebnistabelle nicht abgebildet). 38 % des Varianzanteils an Intelligenz konnten erklärt werden, in erster Linie durch die TEMEKKO Testleistung ($\beta = .43$, $p < .001$, $r^2_{\text{semi-partial}} = .09$). Nächster Prädiktor war interessanterweise das Aggregat aus Gedächtnistests und konnte 4 % einzigartigen Anteil erklären ($\beta = .27$, $p < .01$). Bemerkenswert deshalb, weil der verbliebene Prädiktor Konzentration keinen signifikanten Anteil an der Vorhersage erreichte, wenngleich auch die bivariate Korrelation mit dem Kriterium ($r = .48$) auf dem 1 % Niveau signifikant war. Der TEMEKKO erfasste demnach Fähigkeiten, die zur Vorhersage von Intelligenz besser geeignet waren, als die übrigen in dieser Studie verwendeten Konzentrationstests. Darüber hinaus war zu erkennen, dass intellektuelle

Leistungen zu einem geringen, aber bedeutsamen Anteil auf die Gedächtnisleistungen zurückzuführen waren.

Zusammenfassung der Validitätsschätzungen

Im Rahmen der Prüfung des Strukturgleichungsmodells mit den Konzepten Konzentration, Intelligenz und der TEMEKKO Testleistung wurde für den Einfluss von Materialvarianzanteilen kontrolliert. Hierbei fielen die höchsten korrelativen Zusammenhänge der TEMEKKO Untertests auf das verbale Material ($r = .17$) und auf das numerische Material ($r = .13$), welche als sehr schwache Zusammenhänge galten. Der TEMEKKO wurde als weitgehend unabhängig von Einflüssen einer Inhaltsdomäne eingestuft.

Auf latenter Ebene wies der TEMEKKO einen höheren korrelativen Zusammenhang mit Konzentration ($r = .86$) als mit Intelligenz ($r = .77$) auf, wenn für den Materialeinfluss kontrolliert wurde. Unter Außerachtlassung des Materialeinflusses stieg dieser Zusammenhang auf $r = .90$ mit Konzentration und auf $r = .78$ mit Intelligenz.

Zur Vorhersage des TEMEKKO auf Parcel-Ebene - ohne Berücksichtigung etwaiger Materialeinflüsse - eignete sich demnach erwartungsgemäß Konzentration mit 15 % einzigartig erklärtem Varianzanteil als bester Prädiktor. Die Vorhersagekraft von Intelligenz als zweiter substantieller Prädiktor betrug 7 %. Nahezu der gesamte erklärbare Varianzanteil des TEMEKKO (54 %) konnte auf diese beiden Prädiktoren zurückgeführt werden. Die Gedächtnisleistung konnte in der Vorhersage des TEMEKKO vernachlässigt werden.

Der TEMEKKO eignete sich seinerseits als weitaus bester Prädiktor für Intelligenz, wenngleich seine einzigartige Vorhersagekraft lediglich 9 % am gesamt erklärbaren Varianzanteil von Intelligenz (38 %) betrug. Dennoch war der TEMEKKO zur Vorhersage von Intelligenz besser geeignet als die übrigen Konzentrationstests. Der Prädiktion von Intelligenz kam in dieser Analyse der Gedächtnisleistung mit einem einzigartigen Anteil von 4 % eine größere Bedeutung zu, als bei der Vorhersage des TEMEKKO.

Validierungsstudie 4

Der Datensatz, der als Validierungsstudie 4 verwendet wurde, stammt aus einer Untersuchung einer Diplomandin der Philipps-Universität Marburg, Fachbereich Psychologie.

1 Testdurchführung

Die Erhebung der Daten fand von September bis Dezember 2004 im diagnostischen Labor des Psychologischen Instituts der Philipps-Universität Marburg statt. Die Untersuchungen wurden als Gruppentestungen mit je drei bis fünf Teilnehmern von jeweils einem der drei zur Verfügung stehenden Versuchsleiter durchgeführt.

Die Testung der Probanden wurde auf drei Sitzungen zu je ca. drei Stunden aufgeteilt. Zu Beginn der ersten Testsitzung wurden das Ziel und der Umfang der Untersuchung dargelegt sowie die Abiturzeugnisse der Teilnehmer anonymisiert und kopiert. Am Ende der dritten Testsitzung wurden den Teilnehmern die zugesagten Versuchspersonenstunden auf ihren Karten bestätigt und sie erhielten auf Wunsch Rückmeldungen zu zwei der eingesetzten Verfahren schriftlich und in standardisierter Form. Die Rückmeldung beinhaltete eine Erklärung zum Zweck des jeweiligen Verfahrens und eine Interpretations- bzw. Verständnishilfe zu den Ergebnissen. Um Ermüdungseffekte gering zu halten wurden in den ersten beiden Testsitzungen nach ca. 1 ½ Stunden der Untersuchung eine kurze Pause gehalten. Die dritte Sitzung beinhaltete zwei Pausen.

2 Stichprobe

Der Datensatz umfasste 121 Personen im Alter von 18 bis 46 Jahren ($M = 21.1$ Jahre; $SD = 3.25$ Jahre). Das mittlere Alter der 97 weiblichen Versuchspersonen (80.2 %) betrug 20.8 Jahre ($SD = 2.41$ Jahre), das der 24 männlichen Probanden 22.4. Jahre ($SD = 5.37$ Jahre).

Alle Teilnehmer dieser Untersuchung waren Studierende im Studienfach Psychologie an der Philipps-Universität Marburg, 91.7 % ($N = 111$) im Hauptfach und 8.3 % ($N = 10$) im Nebenfach. 24 Personen dieser Stichprobe hatten zuvor bereits einen anderen Studiengang belegt, 3 Personen hiervon diesen auch abgeschlossen.

Die Anwerbung der Teilnehmer erfolgte größtenteils bei der Immatrikulation zum Studium im Studentensekretariat der Philipps-Universität Marburg, aber auch über Durchsagen in Vorlesungen, Aushänge an schwarzen Brettern des psychologischen Instituts und über Mund-zu-Mund-Propaganda.

Als Entlohnung für die Teilnahme an der Testung erhielten die Studenten Versuchspersonenstunden gutgeschrieben sowie auf Wunsch eine schriftliche, standardisierte Rückmeldung ihrer Testergebnisse des I-S-T 2000 R und des NEO-PI-R.

3 Versuchsmaterial

Der überwiegende Teil der durchgeführten Verfahren wurde für die vorliegende Validierungsstudie nicht berücksichtigt, i.e. MMG, OLMT, LMI, MAE, MultiFlux, NEO-

PI-R, Temint und Stroop. Sie wurden für andere Forschungen erhoben und blieben hier unberücksichtigt. Nach Angaben der jeweiligen Testautoren wurden die Variablen des zu validierenden Verfahrens TEMEKKO als Maß für Konzentration angesehen, die Leistungsmaße der Tests d2 und des FAKT galten als Indikatoren für Konzentration/Aufmerksamkeit. Die Werte aus dem I-S-T 2000 R erfassten Intelligenz und die Subtests der Testbatterie von Oberauer, Süß, Wilhelm und Wittmann (2003) stellten Maße der Arbeitsgedächtniskapazität dar.

Tabelle IV.1: Ablaufplan der Testung

Testverfahren	Untertest	Dauer	Art der Bearbeitung
Eingangsfragebogen			Paper-Pencil
<i>Testsitzung I</i>			
MMG		10 min	PC-gestützt
OLMT		20 min	PC-gestützt
LMI		10 min	PC-gestützt
I-S-T 2000 R		140 min	
	Satzergänzung (verbal)		PC-gestützt
	Analogien (verbal)		PC-gestützt
	Gemeinsamkeiten (verbal)		PC-gestützt
	Rechenaufgaben (num)		PC-gestützt
	Zahlenreihen (num)		PC-gestützt
	Rechenzeichen (num)		PC-gestützt
	Figurenauswahl (fig)		PC-gestützt
	Würfelaufgaben (fig)		PC-gestützt
	Matrizen (fig)		PC-gestützt
<i>Testsitzung II</i>			
Test d2		15 min	Paper-Pencil
TEMEKKO		25 min	Paper-Pencil
MAE		10 min	PC-gestützt
MultiFlux		45 min	PC-gestützt
FAKT		10 min	PC-gestützt
NEO-PI-R		40 min	PC-gestützt
Temint		20 min	PC-gestützt
<i>Testsitzung III</i>			
Stroop-Test		15 min	PC-gestützt
Arbeitsgedächtnis- Testbatterie		135 min	PC-gestützt

3.1 TEMEKKO

Für eine detaillierte Beschreibung des Verfahrens sei auf Abschnitt 3 (Entwicklung des TEMEKKO) im Manual verwiesen.

3.2 Intelligenz-Struktur-Test 2000, revidierte Fassung (I-S-T 2000 R)

(Amthauer, Brocke, Liepmann & Beauducel, 2001). Für eine detaillierte Beschreibung des Verfahrens und seiner Untertests sei auf Abschnitt 3.4 (Intelligenz-Struktur-Test 2000, revidierte Fassung) in Anhang II (Validierungsstudie 2) verwiesen.

3.3 Aufmerksamkeits-Belastungs-Test (d2)

(Brickenkamp, 1987). Die detaillierte Beschreibung des Verfahrens findet sich in Abschnitt 3.2 (Aufmerksamkeits-Belastungs-Test) in Anhang I (Validierungsstudie 1).

3.4 Frankfurter Adaptiver Konzentrationsleistungs-Test (FAKT)

(Moosbrugger & Heyden, 1998). Das Verfahren dient der vollkommen computergestützten Erfassung der individuellen Konzentrationsleistung. Für die vorliegende Untersuchung wurde die Testform SR eingesetzt. Diese verzichtet auf die Adaptivität des Testtempos und erhebt die Reaktionszeiten für die Itembearbeitung, was einer niedrigen situativen Belastung entspricht. Die Aufgabe besteht aus x Reihen zu je 10 Items. Ein Item besteht

aus einem Kreis, in dem entweder noch ein kleinerer Kreis oder ein Quadrat abgebildet sind. Innerhalb dieser im Kreis liegenden Figuren befinden sich zwei oder drei Punkte, die in unterschiedlicher Anordnung stehen, was für die Itembeantwortung aber irrelevant ist. Als Targets sind Kreise mit drei Punkten und Quadrate mit zwei Punkten definiert. Das aktuell vom Probanden zu bearbeitende Item wird vom Programm mit einem Pfeil angezeigt. Ist dieses Item ein Target, ist so schnell als möglich eine „1“ zu drücken. Handelt es sich um einen Distraktor, ist mit „0“ zu antworten. Als Leistungsmaß wurde die Anzahl der richtig beantworteten Items gewertet.

3.5 Arbeitsgedächtnis-Testbatterie (2003)

Die Testbatterie zur Erfassung der Arbeitsgedächtniskapazität von Oberauer, Süß, Wilhelm und Wittmann (2003) besteht aus insgesamt 20 Einzeltests. Sie lassen sich in vier unterschiedliche Aufgabentypen unterteilen und erfassen die drei funktionalen Komponenten des Arbeitsgedächtnisses, wie sie von Oberauer et al. (2003) vorgeschlagen wurden; i.e. Supervision, Speichern im Kontext der Verarbeitung und Koordination.

3.5.1 Wahlreaktionsaufgaben

Dieser Aufgabentyp besteht aus acht binären Wahlreaktionsaufgaben (CRT = Choice Reaction Tasks). Jeweils zwei Aufgaben verwenden das gleiche Stimulus-Material (verbal: Worte, numerisch: Zahlen, figural: Pfeile oder Muster). In der Mitte des Bildschirms erscheint in einem Quadrat der Stimulus, der bis zur Reaktion sichtbar bleibt. Der Proband muss so schnell und korrekt wie möglich das Item beantworten. Für ein verbales Item

beispielsweise musste durch Tastendruck von „<“, das für „links“ bzw. „Tier“ stand oder mit „x“, das für „rechts“ oder „Pflanze“ stand, geantwortet werden. Ein Übungsblock mit 16 Items wurde jeweils vor den fünf Testblöcken zu ebenfalls je 16 Items durchgeführt. Als Leistungsmaß wurden die Reaktionszeit und die Richtigkeit der Zuordnung erfasst.

Die Erfassung der CRTs diene als Baseline zur Ermittlung der generellen Switching Kosten. Diese wurden zusammen mit den Werten der Switching Aufgaben (vgl. 3.5.2 Switching-Aufgaben) als Kennwerte des Komponente Supervision verrechnet. Nachfolgende Tabelle gibt einen Überblick über die einzelnen CRT Subtests.

Tabelle IV.2: Kurzbeschreibung der CRT-Subtests

CRT Subtest	Stimulus-Material	Entscheidungskriterium
verbal: Kategorien	Tier- oder Pflanzenbezeichnungen	Gehört das Wort zur Kategorie <i>Tiere</i> oder Kategorie <i>Pflanzen</i> ?
verbal: Silben	Wörter	Einsilbiges oder zweisilbiges Wort?
numerisch: gerade/ungerade	dreistellige Zahlen	Gerade oder ungerade Zahl?
numerisch: groß/klein	dreistellige Zahlen	Zahl kleiner oder größer als 500?
figural: Pfeile oben/unten	Pfeile	Pfeil in der oberen oder in der unteren Hälfte des Quadrats?
figural: Pfeile auf- /abwärts	Pfeile	Pfeil nach oben oder nach unten gerichtet?
figural: Muster Teile	3x3 Matrix mit einigen geschwärzten Feldern	Bilden die geschwärzten Felder eine zusammenhängende Fläche oder zwei separate Flächen?
figural: Muster Symmetrie	3x3 Matrix mit einigen geschwärzten Feldern	Ist das ausgefüllte Muster entlang einer senkrechten oder waagrechten Achse symmetrisch oder nicht?

3.5.2 Switching-Aufgaben

Die Stimuli dieser insgesamt vier Aufgaben erscheinen nacheinander im Uhrzeigersinn in einer 2 x 2 Matrix auf dem Bildschirm. Für die oberen beiden Felder des Quadrats gilt eine andere Entscheidungsregel (z.B. numerisch: gerade oder ungerade Zahl) als für die beiden unteren Felder (z.B. numerisch: größer oder kleiner als 500). Demnach gelten für 50 % der dargebotenen Stimuli eine andere Entscheidungsregel als im vorangegangenen Item; diese Items gelten als Switching-Items. Die Items, bei denen die gleiche Entscheidungsregel anzuwenden ist wie bei dem vorangegangenen Item, sind Nicht-Switching-Items. Vor den vier Testblöcken zu je 16 Items (1 verbal, 1 numerisch, 2 figural) wird ein Übungsblock mit ebenfalls 16 Items durchgeführt.

Die Switching Aufgaben dienen zusammen mit den Wahlreaktionsaufgaben (vgl. 3.5.1 Wahlreaktionsaufgaben) zur Verrechnung der Komponente Supervision. Die Switching Aufgaben werden ebenso den verschiedenen Materialarten zugeordnet; i.e. switching verbal, switching numerisch und switching figural (bestehend aus Switching Pfeile und Switching Muster). Die Wahlreaktionsaufgaben und Switching Aufgaben, die jeweils die selbe Materialmodalität verwenden, werden miteinander verrechnet. Aus den Switching Aufgaben lassen sich zum einen die Werte der Switching-Items und zum anderen die Werte der Nicht-Switching-Items verwenden. Diese werden entweder zur Berechnung von spezifischen oder von generellen Switching Kosten verwendet. Generelle Switching Kosten sind definiert als die Differenz zwischen log-transformierten Nicht-Switching Reaktionszeiten und dem Mittelwert der beiden inhaltsanalogen Wahlreaktionsaufgaben. Für die vorliegende Studie wurden diese generellen Switching-Kosten als Analysevariable verwendet.

3.5.3 Doppelaufgaben

Bei diesen vier Doppelaufgaben (Dual Tasks) werden Reaktionszeitaufgaben mit Kurzzeitgedächtnis-Aufgaben der gleichen Inhaltsdomäne kombiniert. Zunächst wird das zu erinnernde Material nacheinander mit einem Inter-Stimulus-Intervall von einer Sekunde am Bildschirm dargeboten. Danach sind eine Reihe von Wahlreaktionsaufgaben (vgl. 3.5.1 Wahlreaktionsaufgaben) zu bearbeiten. Die Dauer der Bearbeitung der Wahlreaktionsaufgaben ist auf fünf Sekunden beschränkt, unabhängig davon, wie viele dieser Aufgaben bewältigt werden können. So kann sichergestellt werden, dass das Zeitintervall zwischen Lernen und Abruf bei jeder Doppelaufgabe konstant gehalten wird. Danach werden die memorierten Stimuli in richtiger Reihenfolge wiedergegeben. Die vier unterschiedlichen Doppelaufgaben sind an die jeweiligen Materialien gebunden; i.e. verbal: Wörter, numerisch: Zahlen, figural: Punkte oder Muster.

Die Doppelaufgaben werden miteinander zur Komponente Speichern im Kontext der Verarbeitung verrechnet. Beispielhaft sei die figurale Aufgabe Punkte nachfolgend beschrieben. Innerhalb einer 10 x 10 Matrix auf dem Bildschirm erscheinen nacheinander in den verschiedenen Feldern Punkte. Im darauf folgenden Inter-Stimulus-Intervall (5 Sekunden) sind Wahlreaktionsaufgaben der selben inhaltlichen Domäne zu bearbeiten. Danach sind die Felder, an denen die Punkte erschienen waren, in richtiger Reihenfolge anzuklicken. Die Anzahl der Punkte variierte von zwei bis vier. Als Leistungsmaß wurden die Anzahl der richtig erinnerten Elemente als Gedächtnis-Leistung sowie die logarithmierten Reaktionszeiten der Wahlreaktionsaufgaben erfasst. Tatsächlich weiter verwendet wurden lediglich die Werte der Gedächtnis-Leistung. Begründet wird dies durch Befunde von Oberauer, Süß, Wilhelm und Wittmann (2003) über üblicherweise geringe Korrelationen dieser beiden Untertests. Darüber hinaus ist es Usus, als Indikatoren für

Speichern und Verarbeiten lediglich die Gedächtnis-Leistung zu verwenden (z. B. Daneman & Carpenter, 1980).

3.5.4 Überwachungsaufgaben

Die vier Überwachungsaufgaben dienen der Ermittlung der Komponente Koordination und umfassen die Inhaltsbereiche verbal, numerisch und figural (Quadrate finden und Flugüberwachung). Unabhängig von der inhaltlichen Domäne erfassen sie die Fähigkeit zur Koordination mehrerer variierender Elemente. Auf dem Bildschirm befinden sich mehrere Objekte, die sich unabhängig von einander verändern. Diese Objekte sind vom Probanden im Überblick zu behalten, zu überwachen. Sobald eine vorab definierte Beziehung der Objekte zueinander auftritt (z. B. verbal: drei sich reimende Wörter; oder numerisch: drei Zahlen mit den gleichen Endziffern waagrecht, senkrecht oder diagonal), muss so schnell als möglich eine Taste (Leertaste) gedrückt werden. Die zu überwachenden Stimuli sind jederzeit auf dem Bildschirm präsent, sodass die Gedächtnisanforderung quasi ausgeschlossen werden kann.

Beispielhaft sei die figurale Überwachungsaufgabe Quadrate finden geschildert. In einer 10 x 10 Matrix befinden sich je nach Durchgang acht bis zwölf Punkte; zwei davon sind rot. Alle 1.5 Sekunden verändern die roten Punkte ihre Positionen. In jedem Durchgang ergeben sich somit 20 verschiedene Punkte-Konstellationen. Sobald vier Punkte die Ecken eines Quadrats bilden, unabhängig von Position und Größe des Quadrats, ist so schnell als möglich die Leertaste zu drücken. Als Leistungsmaß galten die Treffer abzüglich der Verwechslungsfehler (fälschlicherweise identifizierte kritische Situationen).

4 Statistische Analysen

4.1 Deskriptiva

Nachstehende Tabelle IV.1 zeigt einen Überblick über die Rohdaten der Analysevariablen.

Tabelle IV.3: Deskriptive Statistiken der Testvariablen

Tests	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>	<i>r_{age}</i>	<i>r_{gender}</i>
Te fig	121	19.0	63.0	42.8	7.9	-0.23	0.52	-.15	-.27**
Te num	121	12.0	73.0	40.8	10.7	0.32	0.19	-.06	-.14
Te verb	121	20.0	61.0	42.0	7.5	0.07	0.54	-.09	.07
Te g	121	63.0	183.0	125.5	22.0	0.03	0.42	-.11	-.14
sfd	121	73.0	154.0	114.9	18.4	-0.19	-0.81	-.21*	-.21*
d2	121	106.0	283.0	194.6	33.0	0.15	0.28	-.24**	.11
FAKT	121	159.0	470.0	280.5	57.8	0.34	0.48	-.18*	-.13
dtv	121	1.2	5.0	3.6	0.7	-0.33	0.73	-.12	-.05
dtn	121	2.3	6.0	4.9	0.7	-0.70	0.83	-.05	-.12
dtP	121	0.5	2.7	1.7	0.5	-0.13	-0.47	-.27**	-.16
dtM	121	0.7	2.9	2.1	0.5	-0.89	0.30	-.41**	-.06
moniv	121	-2.2	2.2	1.2	0.7	-1.53	4.85	-.16	-.06
monin	121	-0.9	2.7	1.7	0.5	-1.6	5.51	-.13	.12
Qf	121	-3.7	3.2	1.9	0.8	-3.2	17.24	.03	-.18
Fü	121	-3.8	-0.4	-1.9	0.8	-0.32	-0.61	-.17	-.33**
gswv	121	-0.3	-0.0	-0.1	0.1	-0.70	0.58	-.15	.13
gswn	121	-0.3	0.0	-0.1	0.0	-0.89	2.62	-.26**	.04
gswP	121	-0.5	0.1	-0.1	0.1	-0.71	3.04	.12	.02
gswM	121	-0.1	0.1	-0.0	0.0	0.38	0.25	-.19*	.25**

Anmerkungen: Te fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO gesamt; sfd = Schlussfolgerndes Denken, IST 2000 R; d2 = Test d2; FAKT = Frankfurter Adaptiver Konzentrations-Leistungstest; dtv = dual task verbal, Arbeitsgedächtnis-Testbatterie; dtn = dual task numerisch, Arbeitsgedächtnis-Testbatterie; dtP = dual task figural Punkte, Arbeitsgedächtnis-Testbatterie; dtM = dual task figural Muster, Arbeitsgedächtnis-Testbatterie; moniv = monitoring verbal, Arbeitsgedächtnis-Testbatterie; monin = monitoring numerisch, Arbeitsgedächtnis-Testbatterie; Qf = Quadrate finden figural, Arbeitsgedächtnis-Testbatterie; Fü = Flugüberwachung figural, Arbeitsgedächtnis-Testbatterie; gswv = generelle Switching Kosten verbal, Arbeitsgedächtnis-Testbatterie; gswn = generelle Switching Kosten numerisch, Arbeitsgedächtnis-Testbatterie; gswP = generelle Switching Kosten figural, Pfeile, Arbeitsgedächtnis-Testbatterie; gswM = generelle Switching Kosten figural, Muster, Arbeitsgedächtnis-Testbatterie. r = bivariate Korrelation, * = $p < .05$, ** = $p < .01$.

Die augenscheinliche Inspektion der einzelnen Testvariablen hinsichtlich ihrer Verteilungsformen mittels Histogrammen und zugehörigen Normalverteilungskurven sowie anhand ihrer numerischen Kennwerte für Schiefe und Kurtosis zeigten keine übermäßig außergewöhnlich schiefen oder gewölbten Auffälligkeiten. Lediglich die Monitoring Variable Quadrate finden lag hinsichtlich der Kurtosis (17.24) gemäß der von West, Finch und Curran (1995) postulierten Grenzen für Schiefe < 2 und Kurtosis < 7 außerhalb der Werte. Die kreuzweise angefertigten Streudiagramme gaben nach optischer Überprüfung keinen Anlass, Fälle aus den Auswertungen zu eliminieren. Es lagen zwar einige Werte abseits der breiten Masse, allerdings im Trend. Es wurden daher alle Fälle für die weiteren Berechnungen beibehalten.

4.2 Zusammenhang mit Alter und Geschlecht

Nach z-Standardisierung der Variablen und Umpolung der Reaktionszeiten (Multiplikation mit -1) wurden die Zusammenhänge mit Alter und Geschlecht ermittelt, wobei insbesondere einige substantielle Korrelationen mit Alter aufschienen. Auch in dieser Untersuchung wurde der Einfluss der beiden Variablen Alter und Geschlecht anhand linearer multipler Regressionsanalysen eliminiert. Nachfolgende Tabelle zeigt die korrelativen Zusammenhänge der Analysevariablen vor und nach Auspartialisierung. Alle Korrelationen waren auf dem 1 %o Niveau signifikant. Die optische Prüfung der Zusammenhänge mittels Streudiagrammen zeigte durchgängig erwartungsgemäß einen positiven linearen Trend.

Tabelle IV.4: Zusammenhang z-transformierte Testvariablen und standardisierte Residuen

$z \backslash$ Res	Te fig	Te num	Te verb	Te g	sfd	d2	FAKT	dtv	dtn	dtP	dtM
Te fig	.941										
Te num		.987									
Te verb			.994								
Te g				.980							
sfd					.945						
d2						.970					
FAKT							.968				
dtv								.989			
dtn									.990		
dtP										.939	
dtM											.900

Anmerkungen: TE fig = TEMEKKO figural; Te num = TEMEKKO numerisch; Te verb = TEMEKKO verbal; Te g = TEMEKKO gesamt; sfd = Schlussfolgerndes Denken, IST 2000 R; d2 = Test d; FAKT = Frankfurter Adaptiver Konzentrations-Leistungstest; dtv = dual task verbal, Arbeitsgedächtnis-Testbatterie; dtn = dual task numerisch, Arbeitsgedächtnis-Testbatterie; dtP = dual task figural, Punkte, Arbeitsgedächtnis-Testbatterie; dtM = dual task figural, Muster, Arbeitsgedächtnis-Testbatterie.

$z \backslash$ Res	moniv	monin	Qf	Fü	gswv	gswn	gswP	gswM
moniv	.983							
monin		.988						
Qf			.984					
Fü				.912				
gswv					.983			
gswn						.967		
gswP							.992	
gswM								.957

Anmerkungen: moniv = monitoring verbal, Arbeitsgedächtnis-Testbatterie; monin = monitoring numerisch, Arbeitsgedächtnis-Testbatterie; Qf = Quadrate finden, figural, Arbeitsgedächtnis-Testbatterie; Fü = Flugüberwachung, figural, Arbeitsgedächtnis-Testbatterie; gswv = generelle Switching Kosten verbal, Arbeitsgedächtnis-Testbatterie; gswn = generelle Switching Kosten numerisch, Arbeitsgedächtnis-Testbatterie; gswP = generelle Swtiching Kosten figural, Pfeile, Arbeitsgedächtnis-Testbatterie; gswM = generelle Switching Kosten figural, Muster, Arbeitsgedächtnis-Testbatterie.

4.3 Schätzung fehlender Werte

In dieser Untersuchung lagen keine fehlenden Werte vor, die hätten geschätzt werden müssen.

4.4 Prüfung auf Gleichheit der Testformen und Materialarten

Die zunächst durchgeführte Prüfung der Annahme einer multivariaten Normalverteilung erbrachte keine Bestätigung und machte die Verwendung der Modelltestung unter Einbeziehung der Bollen-Stine-Bootstrap-Korrektur erforderlich. Der Mardia-Test zur Prüfung der Annahme zeigte dies an: (Tempowerte: multivariate Kurtosis = 4.56, c.r. = 2.56, $p = < .001$ bzw. Fehler-Prozent-Werte: multivariate Kurtosis = 21.69, c.r. = 12.18, $p = < .001$). Alle verwendeten Testvariablen (Tempo-Werte sowie auch Fehler-Prozent-Werte) dieser Untersuchung lagen innerhalb der akzeptablen Grenzen hinsichtlich der Schiefe (< 2) und der Kurtosis (< 7) nach West, Finch und Curran (1995).

Die beiden Strukturgleichungsmodelle (Modell a) und b) in Abbildung 3 des Abschnitts 4.1.2 (Prozedere der Datenbearbeitung) im Manual) wiesen sowohl bei der Prüfung auf Passung der Tempowerte (Modell a) ($\chi^2 [16] = 18.93$, $p = \text{n.s.}$, $CFI = .99$, $RMSEA = .04$ [.00; .10], $SRMR = .07$) als auch der Fehler-Prozent-Werte (Modell b) ($\chi^2 [16] = 16.57$, $p = \text{n.s.}$, $CFI = 1.00$, $RMSEA = .02$ [.00; .09], $SRMR = .05$) einen akzeptablen Modellfit auf. Beide Testformen und alle drei Materialarten können demnach als essentiell-parallele Tests anerkannt werden.

Bei alternativer Testung der Materialien (figural, numerisch, verbal) auf deren Unterschiedlichkeit mit Kovarianzanalysen (Alter und Geschlecht als Kovariate) ergaben sich zunächst für den 3-fach Vergleich sowie die Paarung figural : verbal statistisch signifikante Unterschiede, für die die Annahme der Nullhypothese (Gleichheit der Materialien) nicht beibehalten werden konnte. Der Abgleich mit den hypothetisch ermittelten kritischen F -Werten für kleinstmöglich angenommene Effekte unter den strengen Bedingungen wie in den übrigen Validierungsstudien zeigte für alle Vergleiche

statistische Signifikanz und somit auch praktische Bedeutsamkeit. Die Effektgrößen der Mittelwertsdifferenzen für figural : numerisch ($\eta^2 = .02$), numerisch : verbal ($\eta^2 = .02$) sowie des o.g. 3-fach Vergleichs ($\eta^2 = .04$) waren als gering einzustufen. Die Paarung figural : verbal hingegen zeigte eine Effektstärke ($\eta^2 = .08$) im mittleren Bereich. Diese Ergebnisse sprachen für die Bereitstellung von separaten Normen für die verschiedenen Materialarten. Die Ergebnisse der Analysen zeigt nachstehende Tabelle IV.5.

Tabelle IV.5: Prüfung der Materialarten auf Gleichheit

Material	F_{emp}	p	η^2	power'	F_{krit}
figural : numerisch	1.874	.174	.016	.950	0.477
numerisch : verbal	2.485	.118	.021	.975	0.585
figural : verbal	10.118	.002	.079	.983	0.039
figural: numerisch : verbal	4.509	.012	.037	.950	1.067

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert

Die Prüfung der Formen A und B des TEMEKKO auf deren Unterschiedlichkeit zeigte zunächst keine statistischen Signifikanzen für die Vergleiche der Testformen A und B. Nach Abgleich mit den hypothetisch ermittelten kritischen F -Werten für die Prüfung auf Gleichheit unter erschwerten Bedingungen (analog der Berechnungen vorangegangener Validierungsstudien) erlangten die Paarungen figuraler Formen A : B sowie der verbalen Formen A : B statistisch signifikante Werte. Aufgrund der kleinen Effektstärken bei beiden Vergleichen ($\eta^2 = .03$) war die praktische Bedeutung als gering einzustufen. Nichtsdestotrotz veranlassten die Ergebnisse zur Bereitstellung separater Normen für die jeweiligen Testformen A und B. Die Ergebnisse der Analysen hierzu sind in Tabelle IV.6 gelistet.

Tabelle IV.6: Prüfung der Testformen auf Gleichheit

Testformen	F_{emp}	p	η^2	power'	F_{krit}
A : B (figural)	3.082	.082	.025	.975	1.074
A : B (numerisch)	1.935	.167	.016	.950	5.981
A : B (verbal)	3.431	.066	.028	.983	1.699

Anmerkungen: F_{emp} = empirisch ermittelter F -Wert; p = Überschreitungswahrscheinlichkeit; η^2 = Effektgröße partial eta²; power' = angenommene Teststärke von 95 % unter Berücksichtigung der Anpassung nach Bonferroni-Holm; F_{krit} = hypothetisch kritischer F -Wert.

9.1 Reliabilitäten

Die Reliabilitätsschätzungen der in dieser Untersuchung verwendeten Tests werden nachfolgend in Tabelle IV.7 für ihre Tempowerte gelistet.

Tabelle IV.7: Reliabilitätsmaße der verwendeten Verfahren

Test	N	r_{tt}
TEMEKKO figural	121	.85 ^a
TEMEKKO numerisch	121	.92 ^a
TEMEKKO verbal	121	.89 ^a
TEMEKKO gesamt	121	.94 ^a
IST 2000 R ^c , verbale Intelligenz	3.484	.88 ^b
IST 2000 R ^c , numerische Intelligenz	3.484	.95 ^b
IST 2000 R ^c , figurale Intelligenz	3.484	.87 ^b
IST 2000 R ^c , schlussfolgerndes Denken	3.484	.96 ^b
Test d2 ^c	3.176	.97 ^b
FAKT ^c	n.a.	.91 ^b
Arbeitsgedächtnis ^d , Monitoring verbal	133	.76 ^b
Arbeitsgedächtnis ^d , Monitoring numerisch	133	.59 ^b
Arbeitsgedächtnis ^d , Monitoring Quadrate	133	.73 ^b
Arbeitsgedächtnis ^d , Monitoring Flugüberwachung	133	.86 ^b
Arbeitsgedächtnis ^d , Dual Task verbal	133	.89 ^b
Arbeitsgedächtnis ^d , Dual Task numerisch	133	.83 ^b
Arbeitsgedächtnis ^d , Dual Task Punkte	133	.86 ^b
Arbeitsgedächtnis ^d , Dual Task Muster	133	.91 ^b

Tabelle IV.7 wird fortgesetzt

Arbeitsgedächtnis ^d , generelle switching Kosten, verbal	133	.83 ^b
Arbeitsgedächtnis ^d , generelle switching Kosten, numerisch	133	.83 ^b
Arbeitsgedächtnis ^d , generelle switching Kosten, Pfeile	133	.79 ^b
Arbeitsgedächtnis ^d , generelle switching Kosten, Muster	133	.66 ^b

^a Paralleltestreliabilität

^b Cronbach α

^c Reliabilitätsschätzungen aus Testmanual

^d aus Oberauer, Süß, Wilhelm & Wittmann (2003)

4.6 Validität

Tabelle IV.8 gibt einen Überblick über die Zusammenhänge der in dieser Validierungsstichprobe untersuchten Testleistungen.

Hieraus zeigte sich, dass die Testleistungen des TEMEKKO mit jenen des Tests d2 sowie des FAKTs hoch und signifikant korrelierten, was ein Indiz für zu erwartende Schätzprobleme (Kolinearität) bei modellhaften Berechnungen darstellte. Die minderungskorrigierte Korrelation zwischen den Tests d2 und FAKT (als Maße der Konzentration und Aufmerksamkeit) einerseits und dem TEMEKKO andererseits ($r \sim 1$) veranlasste dazu, die Verfahren d2 und FAKT nicht in den Analysen anhand von Strukturgleichungsmodellen zu berücksichtigen.

Ebenso auffällig war die hohe Korrelation zwischen der TEMEKKO Gesamtleistung und schlussfolgerndem Denken als Gesamtwert ($r = .60$). Diesen Zusammenhang galt es - unter anderem - näher zu untersuchen.

Tabelle IV.8: Korrelationsmatrix der Analysevariablen

<i>r</i>	Te f	Te n	Te v	Te g	sfd	d2	FAKT	dtv	dtn	dtP	dtM	mv	mn	Qf	Fü	gsvv	gswN	gswP	gswM
Te f	1	.55**	.46**	.77**	.52**	.51**	.37**	.13	.18*	.32**	.36**	.31**	.44**	.41**	.21*	.08	.01	-.08	-.07
Te n	.55**	1	.67**	.91**	.53**	.62**	.52**	.13	.22*	.35**	.30**	.40**	.39**	.41**	.27**	.15	.21*	-.21*	.05
Te v	.46**	.67**	1	.83**	.47**	.55**	.63**	.27**	.29**	.31**	.24**	.39**	.28**	.42**	.31**	.29**	.27**	-.21*	.11
Te g	.77**	.91**	.83**	1	.60**	.67**	.60**	.21*	.27**	.38**	.35**	.44**	.44**	.49**	.31**	.20*	.20*	-.21*	.04
sfd	.52**	.53**	.47**	.60**	1	.37**	.34**	.29**	.39**	.52**	.49**	.50**	.34**	.48**	.30**	.20*	.06	-.17	.03
d2	.51**	.62**	.55**	.67**	.37**	1	.46**	.07	.21*	.31**	.29**	.42**	.36**	.42**	.22*	.12	.14	-.33**	.07
FAKT	.37**	.52**	.63**	.60**	.34**	.46**	1	.20*	.21*	.30**	.27**	.28**	.42**	.28**	.39**	.17	.18	-.19*	.01
dtv	.13	.13	.27**	.21*	.29**	.07	.20*	1	.56**	.50**	.47**	.12	.04	.16	.21*	.25**	.19*	-.04	.02
dtn	.18*	.22*	.29**	.27**	.39**	.21*	.21*	.56**	1	.49**	.40**	.21*	.09	.18	.26**	.14	.13	-.06	.25**
dtP	.32**	.35**	.31**	.38**	.52**	.31**	.30**	.50**	.49**	1	.73**	.33**	.27**	.26**	.28**	.22*	.18*	-.03	.12
dtM	.36**	.30**	.24**	.35**	.49**	.29**	.28**	.47**	.40**	.73**	1	.30**	.30**	.32**	.24**	.29**	.13	-.04	.07
mv	.31**	.40**	.39**	.44**	.50**	.42**	.28**	.12	.21*	.33**	.30**	1	.48**	.59**	.11	.07	-.01	-.01	.19*
mn	.44**	.39**	.28**	.44**	.34**	.36**	.42**	.04	.09	.27**	.30**	.48**	1	.41**	.13	.08	-.05	.00	.08
Qf	.41**	.41**	.42**	.49**	.48**	.42**	.28**	.16	.18	.26**	.32**	.59**	.41**	1	.21*	.23*	.10	-.04	.12
Fü	.21*	.27**	.31**	.31**	.30**	.22*	.39**	.21*	.26**	.28**	.34**	.11	.13	.21*	1	.18*	.15	-.09	-.02
gsv	.08	.15	.29**	.20*	.20*	.12	.17	.25**	.14	.22*	.29**	.07	.08	.23*	.18*	1	.56**	.02	.13
gsn	.01	.21*	.27**	.20*	.06	.14	.18	.19*	.13	.18*	.13	-.01	-.05	.10	.15	.56**	1	-.06	.17
gsP	-.08	-.21*	-.21*	-.21*	-.17	-.33**	-.19*	-.04	-.06	-.03	-.04	-.01	.00	-.04	-.09	.02	-.06	1	.13
gsM	-.07	.05	.11	.04	.03	.07	.01	.02	.25**	.12	.07	.19*	.08	.12	-.02	.13	.17	.13	1

Anmerkungen: Te f = TEMEKKO figural; Te n = TEMEKKO numerisch; Te v = TEMEKKO verbal; Te g = TEMEKKO gesamt; sfd = schlussfolgerndes Denken, IST 2000 R; d2 = Test d2; FAKT = Frankfurter Adaptiver Konzentrations-Leistungstest; dtv = dual task verbal, dtn = dual task numerisch; dtP = dual Task figural Punkte; dtM = dual task figural Muster (Arbeitsgedächtnis-Testbatterie); mv = monitoring verbal; mn = monitoring numerisch; Qf = Quadrate finden figural; Fü = Flugüberwachung figural (Arbeitsgedächtnis-Testbatterie); gsv = generelle Switching Kosten verbal; gsn = generelle Switching Kosten numerisch; gsP = generelle Switching Kosten figural, Pfeile; gsM = generelle Switching Kosten figural, Muster (Arbeitsgedächtnis-Testbatterie). *r* = bivariate Korrelation, * = $p < .05$, ** = $p < .01$.

Zunächst wurde ein Messmodell generiert, das die in dieser Studie analysierten Konzepte zu einander in Beziehung setzte. Die Komponenten der Arbeitsgedächtniskapazität stellten die latenten Faktoren Supervision, Coordination und Speichern im Kontext gleichzeitiger Verarbeitung dar, die Skalen des I-S-T 2000 R bildeten den Intelligenz-Faktor und die Untertests des TEMEKKO galten als Indikatoren für Konzentration. Für die wesentlichen Materialanteile wurde kontrolliert, d.h. für das figurale Material wurde mit Ausnahme der Indikatoren für Supervision (generelle switching Kosten) ein Materialfaktor generiert. Die Aussparung der generellen switching Kosten bei figuralen Inhalten erwies sich bei den Entwicklern der Arbeitsgedächtnis-Testbatterie als sinnvoll (Oberauer, Süß, Wilhelm & Wittmann, 2003). Verbalem und numerischem Material wurde lediglich durch die Korrelation zwischen den beiden Dual Task Aufgaben Rechnung getragen.

Die Prüfung des Modells ergab mit $\chi^2 [118] = 154.01$, $p = \text{n.s.}$, $CFI = .95$, $RMSEA = .05$ [.02; .07], $SRMR = .07$ einen akzeptablen Modellfit. Die Konzepte Supervision, Coordination, Speichern im Kontext der Verarbeitung, Intelligenz und des TEMEKKO überschneiden sich zum Teil substantiell ($r = .14$ bis $r = .76$), gleichwohl waren sie aber voneinander abzugrenzen. Die höchsten Zusammenhänge auf latenter Ebene wiesen die TEMEKKO Testleistung und Intelligenz ($r = .76$, $p < .001$), Intelligenz und Coordination ($r = .74$, $p < .001$), Intelligenz und Speichern im Kontext gleichzeitiger Verarbeitung ($r = .69$, $p < .001$) sowie der TEMEKKO und Coordination ($r = .68$, $p < .001$) auf. Die in diesem Modell spezifizierten Parameter zur Schätzung eines Materialfaktors (figural) und eine Korrelation für verbales/numerisches Material bei Dual-Task Aufgaben ($r = .36$) waren ausreichend, um die in den Daten enthaltene Materialvarianz abzubilden. Die figurale Testleistung des TEMEKKO ging maßgeblich in den figuralen Materialfaktor ein ($\lambda = .42$). Dies war bereits aus den Ergebnissen des Vergleichs der Materialarten (vgl. 4.4

Prüfung auf Gleichheit der Testformen und Materialarten) zu erwarten. Das Strukturgleichungsmodell ist in Abbildung IV.1 dargestellt.

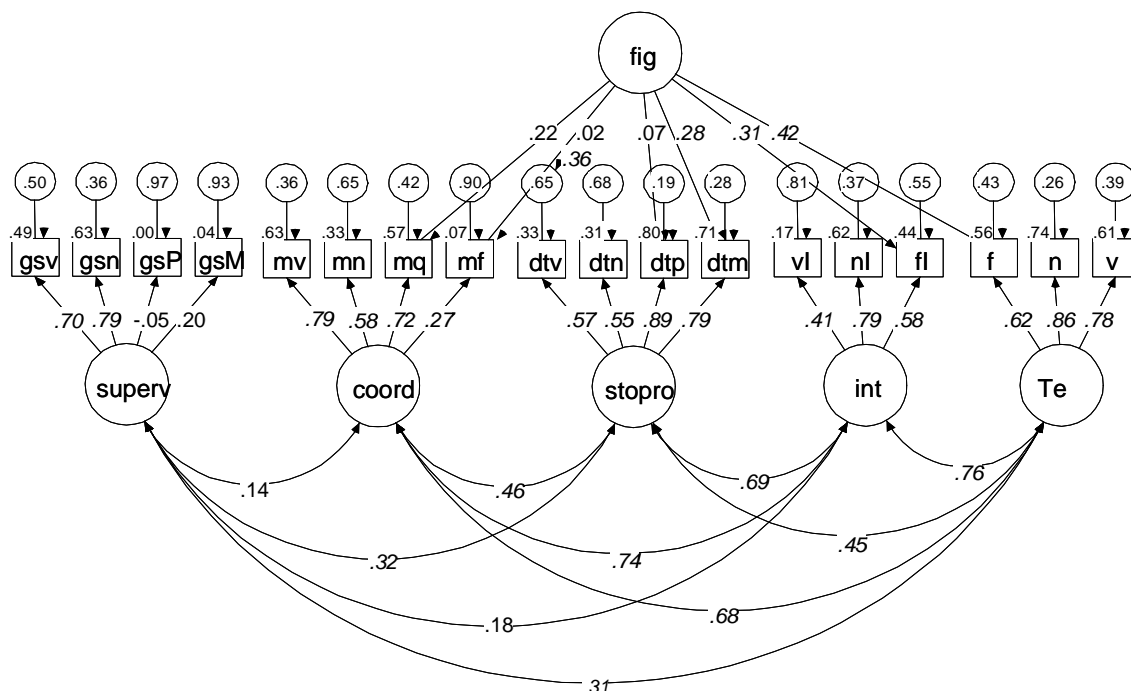


Abbildung IV.1: Modell zur Prüfung der Konstruktstruktur

Anmerkungen: superv = Supervision; coord = Coordination; stopro = Speichern im Kontext gleichzeitiger Verarbeitung; int = Intelligenz; Te = TEMEKKO (Konzentration); fig = figuraler Materialfaktor; gsv = generelle switching Kosten, verbal; gsn = generelle switching Kosten, numerisch; gsP = generelle switching Kosten; Pfeile; gsM = generelle switching Kosten, Muster (aus der Arbeitsgedächtnis-Testbatterie); mv = monitoring, verbal; mn = monitoring, numerisch; mq = monitoring, Quadrate finden; mf = monitoring, Flugüberwachung (aus Arbeitsgedächtnis-Testbatterie); dtv = dual task, verbal; dtn = dual task, numerisch; dtp = dual task, Punkte; dtm = dual task, Muster (aus Arbeitsgedächtnis-Testbatterie); vl = verbale Intelligenz; nl = numerische Intelligenz; fl = figurale Intelligenz (aus IST 2000 R); f = TEMEKKO figural, n = TEMEKKO numerisch, v = TEMEKKO verbal.

Aufbauend auf der Erkenntnis über die zu Grunde liegende Struktur wurde ein um einen g-Faktor auf zweiter Ebene erweitertes Modell gebildet, das in Abbildung IV.2 dargestellt ist. Die Modellprüfung zeigte einen exakten Modellfit ($\chi^2 [123] = 164.58, p = n.s., CFI = .94, RMSEA = .05 [.03; .07], SRMR = .08$). Der g-Faktor konnte aufgrund der hohen Ladung von Intelligenz ($\lambda = .98$) als allgemeine Intelligenz angesehen werden. Der TEMEKKO ($\lambda = .79$) als auch Coordination ($\lambda = .78$) wurden nahe an Intelligenz verortet, wengleich sie auch jeweils unterschiedliche Konzepte erfassten. Alle Ladungen auf dem g-Faktor waren auf dem 1 % Niveau signifikant mit Ausnahme der nicht unbedeutenden Ladung von Supervision ($\lambda = .31, p = .03$).

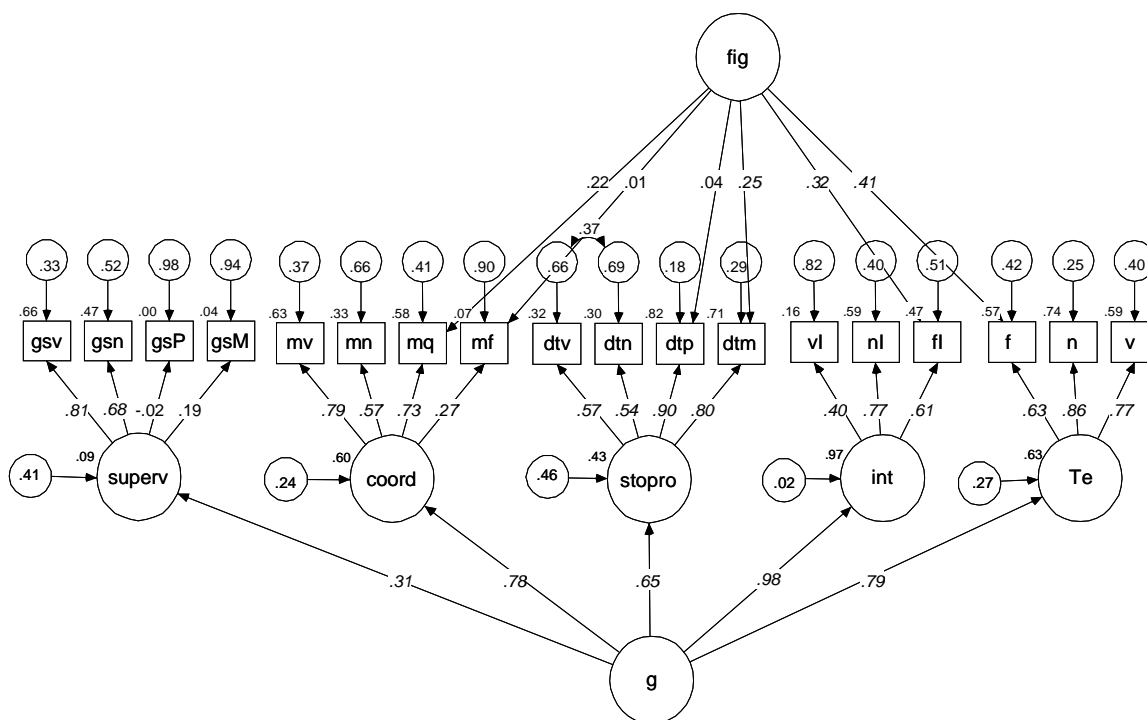


Abbildung IV.2: Modell zur Konstruktstruktur mit g-Faktor

Anmerkungen: superv = Supervision; coord = Coordination; stopro = Speichern im Kontext gleichzeitiger Verarbeitung; int = Intelligenz; Te = TEMEKKO (Konzentration); fig = figuraler Materialfaktor; gsv = generelle switching Kosten, verbal; gsn = generelle switching Kosten, numerisch; gsP = generelle switching Kosten; Pfeile; gsM = generelle switching Kosten, Muster (aus der Arbeitsgedächtnis-Testbatterie); mv = monitoring, verbal; mn = monitoring, numerisch; mq = monitoring, Quadrate finden; mf = monitoring, Flugüberwachung (aus Arbeitsgedächtnis-Testbatterie); dtv = dual task, verbal; dtn = dual task, numerisch; dtp = dual task, Punkte; dtm = dual task, Muster (aus Arbeitsgedächtnis-Testbatterie); vl = verbale Intelligenz; nl = numerische Intelligenz; fl = figurale Intelligenz (aus I-S-T 2000 R); f = TEMEKKO figural, n = TEMEKKO numerisch, v = TEMEKKO verbal.

Unter der Annahme, dass der g-Faktor allgemeine Intelligenz darstellte, galt es zu prüfen, welche Konzepte die beste Vorhersagekraft an der Intelligenz hätten. Hierzu wurden die Analysevariablen einer latenten Regression mit Intelligenz als Kriterium unterzogen. Das Modell zeigte einen exakten Fit, $\chi^2 [118] = 154.01, p = n.s., CFI = .95, RMSEA = .05 [0.02; .07], SRMR = .07$ und ist in Abbildung IV.3 dargestellt. Über alle Konzepte hinweg konnten 79 % des erklärbaren Varianzanteils der Intelligenz vorhergesagt werden.

Die besten Prädiktoren auf latenter Ebene waren die TEMEKKO Testleistung ($\beta = .42, p < .01$) und Speichern im Kontext gleichzeitiger Verarbeitung ($\beta = .42, p < .001$) gleichermaßen. Die Vorhersagekraft von Coordination erreichte knapp keine statistische Signifikanz ($\beta = .28, p = .06.$), obwohl der gemeinsame Varianzanteil mit dem Kriterium

nahezu gleich groß ($r^2_{simple} = .48$) und gleich bedeutsam ($p < .001$) war mit jenem der TEMEKKO Testleistung ($r^2_{simple} = .52$, $p < .001$). Supervision als Prädiktor der Intelligenz ($\beta = -.12$) erwies sich als latenter Suppressor. Wenngleich auch die Bedingungen einer klassischen Suppression oder net-Suppression (Bühner & Ziegler, 2009, S. 685 ff.) nicht vollumfänglich erfüllt waren, so zog Supervision in der Vorhersage von Intelligenz doch Varianz von den übrigen Prädiktoren ab.

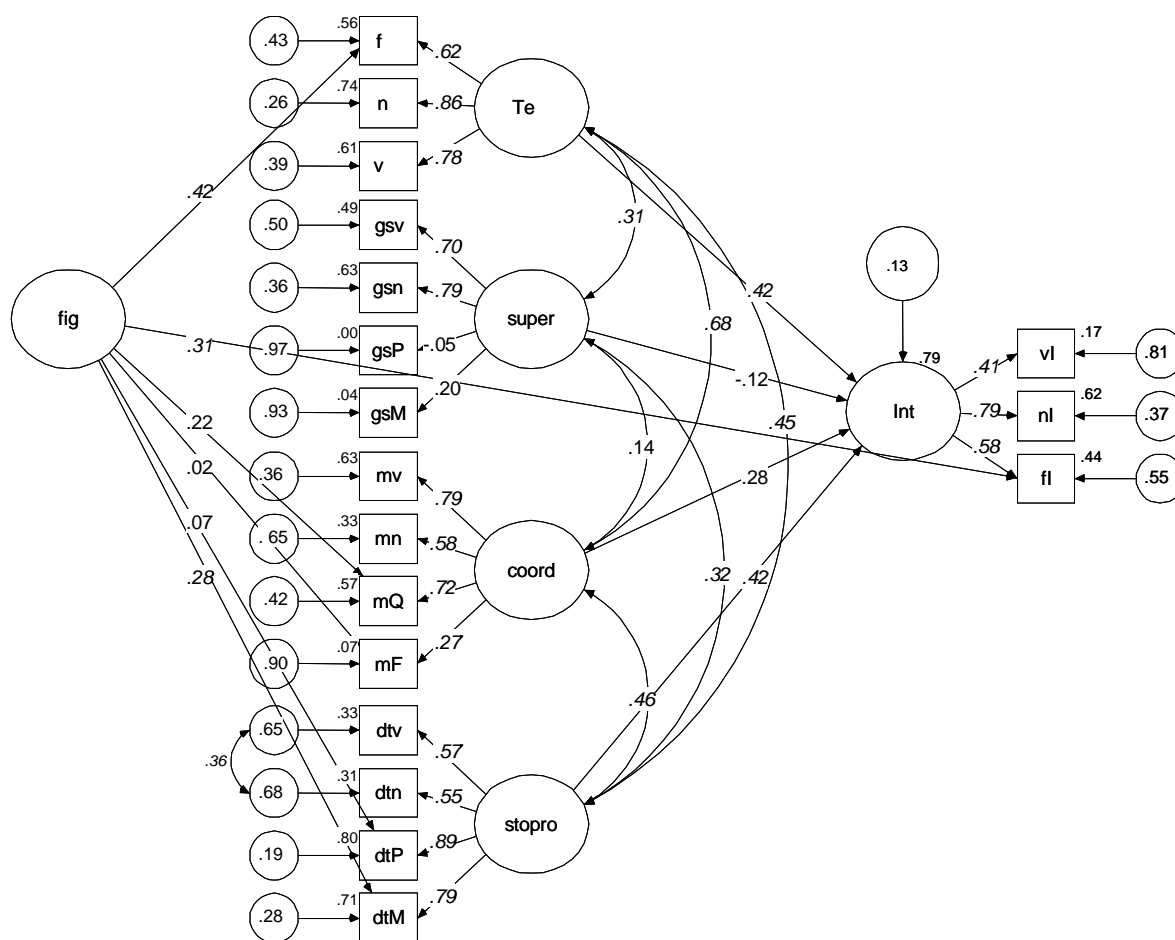


Abbildung IV.3: latente Regression mit Intelligenz als Kriterium

Anmerkungen: superv = Supervision; coord = Coordination; stopro = Speichern im Kontext gleichzeitiger Verarbeitung; int = Intelligenz; Te = TEMEKKO (Konzentration); fig = figurale Materialfaktor; gsv = generelle switching Kosten, verbal; gsn = generelle switching Kosten, numerisch; gsP = generelle switching Kosten; Pfeile; gsM = generelle switching Kosten, Muster (aus der Arbeitsgedächtnis-Testbatterie); mv = monitoring, verbal; mn = monitoring, numerisch; mq = monitoring, Quadrate finden; mf = monitoring, Flugüberwachung (aus Arbeitsgedächtnis-Testbatterie); dtv = dual task, verbal; dtn = dual task, numerisch; dtp = dual task, Punkte; dtm = dual task, Muster (aus Arbeitsgedächtnis-Testbatterie); vl = verbale Intelligenz; nl = numerische Intelligenz; fl = figurale Intelligenz (aus I-S-T 2000 R); f = TEMEKKO figural, n = TEMEKKO numerisch, v = TEMEKKO verbal.

Die Einzelergebnisse der latenten Regression innerhalb des Strukturgleichungsmodells sind in Tabelle IV.9 gesondert gelistet. Die Werte der einzigartigen Vorhersageanteile an dem erklärbaren Varianzanteil der Intelligenz anhand quadrierter Semipartialkorrelationen konnten aus modellhaften Berechnungen (AMOS) nicht ermittelt werden. Die Berechnungsmethode via SPSS wiederum ermöglichte nicht, für Materialeinflüsse zu kontrollieren. Als Maße für die Einzigartigkeit der Vorhersage wurden daher die standardisierten Regressionsgewichte (β -Gewichte) der Berechnung auf latenter Ebene herangezogen.

Tabelle IV.9: Ergebnisse der latenten Regression mit Intelligenz als Kriterium

	β	r^2_{simple}
TEMEKKO	.42 **	.52 ***
Stopro	.42 ***	.37 **
Coordination	.28	.48 ***
Supervision	-.12	.01

AV = Intelligenz, $R^2 = .79$

Anmerkungen: Stopro = Parcel Speichern im Kontext gleichzeitiger Verarbeitung; TEMEKKO = Parcel TEMEKKO; Coordination = Parcel Coordination; Supervision = Parcel Supervision; Intelligenz = Parcel Intelligenz (Zusammensetzung der Parcels wie in Abbildungen IV.1, IV.2 und IV.3). β = standardisierter Regressionskoeffizient, r^2_{simple} = quadrierte bivariate Korrelation, AV = abhängige Variable; R^2 = korrigierter Determinationskoeffizient; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Welche der dieser Untersuchung zugrunde liegenden Konzepte die Testleistungen im TEMEKKO vorhersagen konnten, sollte Aufschluss darüber geben, welche Fähigkeiten bei der Bearbeitung des TEMEKKO erforderlich seien. Dieser Frage wurde ebenso anhand einer latenten Regressionsanalyse unter Kontrolle um das figurale sowie verbal/numerische Material nachgegangen. Hierfür wurde das gleiche Modell verwendet, wie in Abbildung IV.3 dargestellt, nur wurden die latenten Variablen und zugehörigen Indikatoren für Intelligenz und TEMEKKO ausgetauscht. Intelligenz fungierte demnach als Prädiktor, die TEMEKKO Testleistung galt als Kriterium (Modell nicht abgebildet). Das Modell zeigte mit $\chi^2 [118] = 154.01$, $p = \text{n.s.}$, $CFI = .95$, $RMSEA = .05$ [.02; .07], $SRMR = .07$ einen

exakten Fit. 66 % des erklärbaren Varianzanteils der TEMEKKO Testleistung konnten vorhergesagt werden. Der beste Prädiktor auf latenter Ebene war Intelligenz ($\beta = .69$, $p = .03$). Supervision erreichte mit $\beta = .22$ noch knapp Signifikanz ($p = .05$), wohingegen Coordination nicht statistisch bedeutsam zur Vorhersage beitragen konnte ($\beta = .24$, n.s.). Der gemeinsame Varianzanteil von Coordination und dem TEMEKKO aber war auf dem 1 %o Niveau signifikant. Dieser Varianzanteil enthielt offenkundig auch Anteile anderer Prädiktoren, die sich in deren einzigartigen Vorhersagen zeigten. In dieser Regressionsanalyse erwies sich Speichern im Kontext gleichzeitiger Verarbeitung als latenter Suppressor ($\beta = -.21$). Auch hier erfüllte der Prädiktor die geforderten Bedingungen einer klassischen Suppression oder net-Suppression (Bühner & Ziegler, 2009, S. 685 ff.) nicht vollumfänglich. Dennoch schien er zur Vorhersage des TEMEKKO von den übrigen Prädiktoren Varianzanteile abzuziehen.

Tabelle IV.10 zeigt die Einzelergebnisse der latenten Regression innerhalb des Strukturgleichungsmodells. Auch in dieser latenten Regressionsanalyse wurden - anstelle der nicht modellhaft rechenbaren quadrierten Semipartialkorrelationen - die standardisierten Regressionsgewichte (β -Gewichte) als Maß für die Einzigartigkeit der Vorhersage benutzt, um den Materialeinfluss berücksichtigt haben zu können.

Tabelle IV.10: Ergebnisse der latenten Regression mit TEMEKKO als Kriterium

	β	r^2_{simple}
Intelligenz	.69 *	.58 ***
Supervision	.22 *	.10 *
Coordination	.24	.44 ***
Stopro	-.21	.16 *

AV = TEMEKKO, $R^2 = .66$

Anmerkungen: Intelligenz = Parcel Intelligenz; Coordination = Parcel Coordination; Supervision = Parcel Supervision; Stopro = Parcel Speichern im Kontext gleichzeitiger Verarbeitung; TEMEKKO = Parcel TEMEKKO (Zusammensetzung der Parcels wie in Abbildungen IV.1, IV.2 und IV.3). β = standardisierter Regressionskoeffizient, r^2_{simple} = quadrierte bivariate Korrelation; AV = abhängige Variable; R^2 = korrigierter Determinationskoeffizient; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Die zu Beginn der Analysen unterstellte Problematik einer Kolinearität zwischen dem TEMEKKO und den Maßen d2 bzw. FAKT wurde abschließend nochmals aufgegriffen. Die beiden Analysevariablen d2 bzw. FAKT flossen auf Parcel-Ebene als Prädiktor in eine lineare multiple Regressionsanalyse ein, bei der allerdings nicht für Materialvarianzanteile kontrolliert werden konnte. Gleichwohl machten die Ergebnisse deutlich, wie eng diese Maße mit der TEMEKKO Testleistung verknüpft sind. Das Parcel dieser beiden Tests war der deutlich beste Prädiktor ($\beta = .58$) in der Vorhersage des erklärten Varianzanteils (65 %) des TEMEKKO. Es konnte 21 % dieses Anteils erklären. Einen ebenso nicht unerheblichen Beitrag zur Vorhersage leistete - wie aus der Analyse unter Berücksichtigung der Materialvarianzanteile ersichtlich - Intelligenz (6 %).

Tabelle IV. 11: Ergebnisse der Regressionsanalyse mit dem TEMEKKO als Kriterium auf Parcel-Ebene

	β	r^2_{simple}	$r^2_{semi-partial}$
Konzentration	.58 ***	.56 ***	.21 ***
Intelligenz	.33 ***	.36 ***	.06 ***
Coordination	.08	.36 ***	< .01
Supervision	.05	.01	.002
Stopro	-.02	.15 ***	< .001

AV = TEMEKKO, $R^2 = .65$

Anmerkungen: Konzentration = Parcel Konzentration (d2, FAKT); Intelligenz = Parcel Intelligenz, Coordination = Parcel Coordination, Supervision = Parcel Supervision, Stopro = Parcel Speichern im Kontext gleichzeitiger Verarbeitung (Zusammensetzung der Parcels wie in Abbildungen IV.1, IV.2 und IV.3). β = standardisierter Regressionskoeffizient, r^2_{simple} = quadrierte bivariate Korrelation, $r^2_{semi-partial}$ = quadrierte Semipartialkorrelation, AV = abhängige Variable; R^2 = korrigierter Determinationskoeffizient; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Zusammenfassung der Validitätsschätzungen

Die Testleistungen aus dem TEMEKKO einerseits und jene aus den Tests d2 und FAKT, als Maße der Konzentration und Aufmerksamkeit erwiesen sich als nicht weiter separierbar

(minderungskorrigierte Korrelation $r \sim 1$). Die Vorhersage des erklärbaren Varianzanteils des TEMEKKO (65 %) auf Parcel-Ebene unter Berücksichtigung der beiden Verfahren zeigte das Aggregat d2/FAKT demnach erwartungsgemäß als den weitaus besten Prädiktor (21 % einzigartig erklärter Varianzanteil).

Unter Außerachtlassung der Tests d2 und FAKT, aber mit Berücksichtigung hinreichend kontrollierter Materialvarianzanteile zeigten die analysierten Konzepte - Intelligenz, Coordination, Speichern im Kontext gleichzeitiger Verarbeitung, Supervision und die TEMEKKO Testleistung - auf latenter Ebene substantielle Überschneidungsbereiche ($r = .14$ bis $r = .76$ minderungskorrigiert), insbesondere zwischen TEMEKKO und Intelligenz ($r = .76$), Intelligenz und Coordination ($r = .74$), Intelligenz und Speichern im Kontext gleichzeitiger Verarbeitung ($r = .69$) sowie dem TEMEKKO und Coordination ($r = .68$). Nichtsdestotrotz handelte es sich um separierbare Konzepte.

Wurde für einen übergeordneten Faktor auf zweiter Ebene kontrolliert, so betrug die Ladung von Intelligenz auf diesem g -Faktor $\lambda = .98$. Demnach wurde dieser Faktor zweiter Ordnung als allgemeine Intelligenz verstanden. Die TEMEKKO Testleistung sowie Coordination ließen sich mit $\lambda = .79$ bzw. $\lambda = .78$ dort verorten. Diese hohen Zusammenhänge wiesen auf die konzeptionelle Nähe zwischen den Konstrukten hin.

Zur Vorhersage des erklärbaren Varianzanteils von Intelligenz (79 %) auf latenter Ebene eigneten sich der TEMEKKO und Speichern im Kontext gleichzeitiger Verarbeitung als dessen stärkste Prädiktoren ($\beta = .42$) jeweils gleich gut. Coordination verzeichnete mit 48 % einen substantiellen gemeinsamen Varianzanteil mit Intelligenz, war aber für dessen Vorhersage von untergeordneter Bedeutung. Der erklärbare Varianzanteil der TEMEKKO Testleistung auf latenter Ebene betrug 66 %. Als maßgebliche Prädiktoren hierbei erwiesen

sich Intelligenz ($\beta = .69$) und Supervision ($\beta = .22$). Auch in dieser Untersuchung teilte Coordination mit dem Kriterium wesentliche Varianzanteile (44 %). Zur Vorhersage einzigartiger Anteile am TEMEKKO allerdings waren sie aus statistischer Sicht nicht relevant.

Tabelle V. 1: Übersicht über Verteilung der Targets und Distraktoren in Testform A sowie über die Platzierung besonders leichter und besonders schwerer Items

Testform A

	01	02	03	04	05	06	07	08	09	10
01-10	T	D	d0	d	D	D	to	D	D	T
11-20	D	T	D	D	D	D	D	D	T	D
21-30	D	D	D	to	T	D	d	D	T	D
31-40	D	d0	D	D	T	to	D	D	D	D
41-50	D	T	d0	D	T	D	D	d0	D	d0
51-60	D	t	D	D	d0	D	D	T	T	D
61-70	D	d0	D	T	D	D	D	d0	T	d
71-80	d0	T	D	D	D	D	T	D	D	T
81-90	D	d0	D	T	D	d	T	d0	D	D
91-100	D	D	T	D	D	D	T	D	d0	T
101-110	d0	d	to	D	D	T	D	d0	D	D
111-120	d0	D	D	d	T	T	D	D	to	D
121-130	D	D	D	T	d0	D	d	T	T	D
131-140	T	d	D	D	T	D	D	D	D	D
141-150	D	T	D	D	d0	D	D	t	D	T
151-160	D	D	D	D	T	T	D	D	d	T
161-170	D	D	D	d0	D	T	D	T	D	D
171-180	D	D	T	D	D	T	D	D	to	d0
181-190	d	D	to	d0	D	D	D	T	D	D

T besonders schweres Target
t besonders leichtes Target
to besonders leichtes Target, ohne Zusatzmerkmale
T durchschnittliches Target

D besonders schwerer Distraktor
d besonders leichter Distraktor
d0 besonders leichter Distraktor, ohne Zusatzmerkmale
D durchschnittlicher Distraktor

Tabelle V. 2: Übersicht über Verteilung der Targets und Distraktoren in Testform B sowie über die Platzierung besonders leichter und besonders schwerer Items

Testform B

	01	02	03	04	05	06	07	08	09	10
01-10	d0	d	to	D	D	T	D	d0	D	D
11-20	d0	D	D	d	T	T	D	D	to	D
21-30	D	D	D	T	d0	D	d	T	T	D
31-40	T	d	D	D	T	D	D	D	D	D
41-50	D	T	D	D	d0	D	D	t	D	T
51-60	D	D	D	D	T	T	D	D	d	T
61-70	D	D	D	d0	D	T	D	T	D	D
71-80	D	D	T	D	D	T	D	D	to	d0
81-90	d	D	to	d0	D	D	D	T	D	D
91-100	T	D	d0	d	D	D	to	D	D	T
101-110	D	T	D	D	D	D	D	D	T	D
111-120	D	D	D	to	T	D	d	D	T	D
121-130	D	d0	D	D	T	to	D	D	D	D
131-140	D	T	d0	D	T	D	D	d0	D	d0
141-150	D	t	D	D	d0	D	D	T	T	D
151-160	D	d0	D	T	D	D	D	d0	T	d
161-170	d0	T	D	D	D	D	T	D	D	T
171-180	D	d0	D	T	D	d	T	d0	D	D
181-190	D	D	T	D	D	D	T	D	d0	T

- T** besonders schweres Target
 t besonders leichtes Target
 t0 besonders leichtes Target, ohne Zusatzmerkmale
 T durchschnittliches Target

D besonders schwerer Distraktor
 d besonders leichter Distraktor
 d0 besonders leichter Distraktor, ohne Zusatzmerkmale
 D durchschnittlicher Distraktor

Tabelle V. 3: mögliche Targets

1	3	5
1	3	6
1	3	7
1	3	8
1	3	9
1	4	6
1	4	7
1	4	8
1	4	9
1	5	7
1	5	8
1	5	9
1	6	8
1	6	9
1	7	9
2	4	6
2	4	7
2	4	8
2	4	9
2	5	7
2	5	8
2	5	9
2	6	8
2	6	9
2	7	9
3	5	7
3	5	8
3	5	9
3	6	8
3	6	9
3	7	9
4	6	8
4	6	9
4	7	9
5	7	9

5	3	1
6	4	2
6	4	1
6	3	1
7	5	3
7	5	2
7	5	1
7	4	2
7	4	1
7	3	1
8	6	4
8	6	3
8	6	2
8	6	1
8	5	3
8	5	2
8	5	1
8	4	2
8	4	1
8	3	1
9	7	5
9	7	4
9	7	3
9	7	2
9	7	1
9	6	4
9	6	3
9	6	2
9	6	1
9	5	3
9	5	2
9	5	1
9	4	2
9	4	1
9	3	1

Tabelle V. 4: mögliche Distraktoren

1	1	1
1	1	3
1	1	4
1	1	5
1	1	6
1	1	7
1	1	8
1	1	9
1	3	1
1	3	3
1	4	1
1	4	2
1	4	4
1	5	1
1	5	2
1	5	3
1	5	5
1	6	1
1	6	2
1	6	3
1	6	4
1	6	6
1	7	1
1	7	2
1	7	3
1	7	4
1	7	5
1	7	7
1	8	1
1	8	2
1	8	3
1	8	4
1	8	5
1	8	6
1	8	8
1	9	1
1	9	2
1	9	3
1	9	4
1	9	5
1	9	6
1	9	7
1	9	9
2	2	2

2	2	4
2	2	5
2	2	6
2	2	7
2	2	8
2	2	9
2	4	1
2	4	2
2	4	4
2	5	1
2	5	2
2	5	3
2	5	5
2	6	1
2	6	2
2	6	3
2	6	4
2	6	6
2	7	1
2	7	2
2	7	3
2	7	4
2	7	5
2	7	7
2	8	1
2	8	2
2	8	3
2	8	4
2	8	5
2	8	6
2	8	8
2	9	1
2	9	2
2	9	3
2	9	4
2	9	5
2	9	6
2	9	7
2	9	9
3	1	1
3	1	3
3	1	4
3	1	5
3	1	6

3	1	7
3	1	8
3	1	9
3	3	1
3	3	3
3	3	5
3	3	6
3	3	7
3	3	8
3	3	9
3	5	1
3	5	2
3	5	3
3	5	5
3	6	1
3	6	2
3	6	3
3	6	4
3	6	6
3	7	1
3	7	2
3	7	3
3	7	4
3	7	5
3	7	7
3	8	1
3	8	2
3	8	3
3	8	4
3	8	5
3	8	6
3	8	8
3	9	1
3	9	2
3	9	3
3	9	4
3	9	5
3	9	6
3	9	7
3	9	9
4	1	1
4	1	3
4	1	4
4	1	5

4	1	6
4	1	7
4	1	8
4	1	9
4	2	2
4	2	4
4	2	5
4	2	6
4	2	7
4	2	8
4	2	9
4	4	1
4	4	2
4	4	4
4	4	6
4	4	7
4	4	8
4	4	9
4	6	1
4	6	2
4	6	3
4	6	4
4	6	6
4	7	1
4	7	2
4	7	3
4	7	4
4	7	5
4	7	7
4	8	1
4	8	2
4	8	3
4	8	4
4	8	5
4	8	6
4	8	8
4	9	1
4	9	2
4	9	3
4	9	4
4	9	5
4	9	6
4	9	7
4	9	9

Tabelle V.4 wird fortgesetzt

5	1	1
5	1	3
5	1	4
5	1	5
5	1	6
5	1	7
5	1	8
5	1	9
5	2	2
5	2	4
5	2	5
5	2	6
5	2	7
5	2	8
5	2	9
5	3	3
5	3	5
5	3	6
5	3	7
5	3	8
5	3	9
5	5	1
5	5	2
5	5	3
5	5	5
5	5	7
5	5	8
5	5	9
5	7	1
5	7	2
5	7	3
5	7	4
5	7	5
5	7	7
6	1	1
6	1	3
6	1	4
6	1	5
6	1	6
6	1	7
6	1	8
6	1	9
6	2	2
6	2	4

6	2	5
6	2	6
6	2	7
6	2	8
6	2	9
6	3	3
6	3	5
6	3	6
6	3	7
6	3	8
6	3	9
6	4	4
6	4	6
6	4	7
6	4	8
6	4	9
6	6	1
6	6	2
6	6	3
6	6	4
6	6	6
6	6	8
6	6	9
6	8	1
6	8	2
6	8	3
6	8	4
6	8	5
6	8	6
6	8	8
6	9	1
6	9	2
6	9	3
6	9	4
6	9	5
6	9	6
6	9	7
6	9	9
7	1	1
7	1	3
7	1	4
7	1	5
7	1	7
7	1	9

8	1	1
8	1	3
8	1	4
8	1	5
8	1	6
8	1	7
8	1	8
8	1	9
8	2	2
8	2	4
8	2	5
8	2	6
8	2	7
8	2	8
8	2	9
8	3	3
8	3	5
8	3	6
8	3	7
8	3	8
8	3	9
8	4	4
8	4	6
8	4	7
8	4	8
8	4	9
8	5	5
8	5	7
8	5	8
8	5	9
8	6	6
8	6	8
8	6	9
8	8	1
8	8	2
8	8	3
8	8	4
8	8	5
8	8	6
8	8	8
9	1	1
9	1	3
9	1	4
9	1	5

9	1	6
9	1	7
9	1	8
9	1	9
9	2	2
9	2	4
9	2	5
9	2	6
9	2	7
9	2	8
9	2	9
9	3	3
9	3	5
9	3	6
9	3	7
9	3	8
9	3	9
9	4	4
9	4	6
9	4	7
9	4	8
9	4	9
9	5	5
9	5	7
9	5	8
9	5	9
9	6	6
9	6	8
9	6	9
9	7	7
9	7	9
9	9	1
9	9	2
9	9	3
9	9	4
9	9	5
9	9	6
9	9	7
9	9	9

Normstichprobe Schüler

$N = 312$

♂: $N = 185$, ♀: $N = 127$

≤ 20 Jahre: $N = 94$, 21 bis 25 Jahre: $N = 194$, 26 bis 30 Jahre: $N = 22$, ≥ 30 Jahre: $N = 2$

KL- Werte	Standardwerte					
	figural		numerisch		verbal	
	Form A	Form B	Form A	Form B	Form A	Form B
1	62	45	69	65	62	60
2	64	48	70	67	65	62
3	66	50	72	69	67	65
4	68	53	73	70	69	67
5	71	55	75	72	71	69
6	73	57	76	74	73	71
7	75	60	78	75	76	73
8	77	62	79	77	78	75
9	79	64	81	79	80	77
10	81	67	83	81	82	79
11	83	69	84	82	85	81
12	86	72	86	84	87	83
13	88	74	87	86	89	85
14	90	76	89	88	91	88
15	92	79	90	89	94	90
16	94	81	92	91	96	92
17	96	83	93	93	98	94
18	98	86	95	95	100	96
19	100	88	96	96	102	98
20	103	91	98	98	105	100
21	105	93	99	100	107	102
22	107	95	101	101	109	104
23	109	98	102	103	111	106
24	111	100	104	105	114	108
25	113	102	105	107	116	111
26	115	105	107	108	118	113
27	118	107	108	110	120	115
28	120	110	110	112	123	117
29	122	112	111	114	125	119
30	124	114	113	115	127	121
31	126	117	114	117	129	123
32	128	119	116	119	131	125
33	130	121	117	121	134	127
34	132	124	119	122	136	129
35	135	126	120	124	138	131
36	137	129	122	126	140	134
37	139	131	123	127	143	136
38	141	133	125	129	145	138
39	143	136	126	131	147	140
40	145	138	128	133	149	142
41	147	140	130	134	151	144
42	150	143	131	136	154	146
43	152	145	133	138	156	148
44	154	148	134	140	158	150
45	156	150	136	141	160	152
46	158	152	137	143	163	155
47	160	155	139	145	165	157
48	162	157	140	147	167	159

Normstichprobe Schüler

 $N = 312$ ♂: $N = 185$, ♀: $N = 127$ ≤ 20 Jahre: $N = 94$, 21 bis 25 Jahre: $N = 194$, 26 bis 30 Jahre: $N = 22$, ≥ 30 Jahre: $N = 2$

KL- Werte	Standardwerte	
	Form A	Form B
1	54	46
2	55	46
3	56	47
4	57	48
5	58	49
6	58	50
7	59	51
8	60	51
9	61	52
10	62	53
11	62	54
12	63	55
13	64	56
14	65	57
15	66	57
16	66	58
17	67	59
18	68	60
19	69	61
20	70	62
21	70	63
22	71	63
23	72	64
24	73	65
25	74	66
26	74	67
27	75	68
28	76	68
29	77	69
30	78	70
31	78	71
32	79	72
33	80	73
34	81	74
35	82	74
36	82	75
37	83	76
38	84	77
39	85	78
40	86	79
41	86	80
42	87	80
43	88	81
44	89	82
45	89	83
46	90	84
47	91	85
48	92	85

KL- Werte	Standardwerte	
	Form A	Form B
49	93	86
50	93	87
51	94	88
52	95	89
53	96	90
54	97	91
55	97	91
56	98	92
57	99	93
58	100	94
59	101	95
60	101	96
61	102	97
62	103	97
63	104	98
64	105	99
65	105	100
66	106	101
67	107	102
68	108	102
69	109	103
70	109	104
71	110	105
72	111	106
73	112	107
74	113	108
75	113	108
76	114	109
77	115	110
78	116	111
79	117	112
80	117	113
81	118	113
82	119	114
83	120	115
84	121	116
85	121	117
86	122	118
87	123	119
88	124	119
89	124	120
90	125	121
91	126	122
92	127	123
93	128	124
94	128	125
95	129	125
96	130	126

KL- Werte	Standardwerte	
	Form A	Form B
97	131	127
98	132	128
99	132	129
100	133	130
101	134	130
102	135	131
103	136	132
104	136	133
105	137	134
106	138	135
107	139	136
108	140	136
109	140	137
110	141	138
111	142	139
112	143	140
113	144	141
114	144	142
115	145	142
116	146	143
117	147	144
118	148	145
119	148	146
120	149	147
121	150	147
122	151	148
123	152	149
124	152	150
125	153	151
126	154	152
127	155	153
128	156	153
129	156	154
130	157	155
131	158	156
132	159	157
133	159	158
134	160	159
135	161	159
136	162	160
137	163	161
138	163	162
139	164	163
140	165	164
141	166	164
142	167	165
143	167	166
144	168	167

Normstichprobe Schüler

$N = 312$

♂: $N = 185$, ♀: $N = 127$

≤ 20 Jahre: $N = 94$, 21 bis 25 Jahre: $N = 194$, 26 bis 30 Jahre: $N = 22$, ≥ 30 Jahre: $N = 2$

KL-Werte	Standardwerte gesamt	KL-Werte	Standardwerte gesamt	KL-Werte	Standardwerte gesamt
1	47	49	68	97	89
2	48	50	69	98	89
3	48	51	69	99	90
4	49	52	69	100	90
5	49	53	70	101	90
6	50	54	70	102	91
7	50	55	71	103	91
8	50	56	71	104	92
9	51	57	72	105	92
10	51	58	72	106	93
11	52	59	72	107	93
12	52	60	73	108	93
13	53	61	73	109	94
14	53	62	74	110	94
15	53	63	74	111	95
16	54	64	75	112	95
17	54	65	75	113	96
18	55	66	75	114	96
19	55	67	76	115	96
20	56	68	76	116	97
21	56	69	77	117	97
22	56	70	77	118	98
23	57	71	78	119	98
24	57	72	78	120	99
25	58	73	78	121	99
26	58	74	79	122	99
27	59	75	79	123	100
28	59	76	80	124	100
29	59	77	80	125	101
30	60	78	81	126	101
31	60	79	81	127	102
32	61	80	81	128	102
33	61	81	82	129	102
34	62	82	82	130	103
35	62	83	83	131	103
36	62	84	83	132	104
37	63	85	84	133	104
38	63	86	84	134	105
39	64	87	84	135	105
40	64	88	85	136	105
41	65	89	85	137	106
42	65	90	86	138	106
43	65	91	86	139	107
44	66	92	87	140	107
45	66	93	87	141	108
46	67	94	87	142	108
47	67	95	88	143	108
48	68	96	88	144	109

Tabelle wird fortgesetzt

Normstichprobe Schüler

$N = 312$

♂: $N = 185$, ♀: $N = 127$

≤ 20 Jahre: $N = 94$, 21 bis 25 Jahre: $N = 194$, 26 bis 30 Jahre: $N = 22$, ≥ 30 Jahre: $N = 2$

KL-Werte	Standardwerte gesamt
145	109
146	110
147	110
148	111
149	111
150	111
151	112
152	112
153	113
154	113
155	114
156	114
157	114
158	115
159	115
160	116
161	116
162	117
163	117
164	117
165	118
166	118
167	119
168	119
169	120
170	120
171	120
172	121
173	121
174	122
175	122
176	123
177	123
178	123
179	124
180	124
181	125
182	125
183	126
184	126
185	126
186	127
187	127
188	128
189	128
190	129
191	129
192	130

KL-Werte	Standardwerte gesamt
193	130
194	130
195	131
196	131
197	132
198	132
199	133
200	133
201	133
202	134
203	134
204	135
205	135
206	136
207	136
208	136
209	137
210	137
211	138
212	138
213	139
214	139
215	139
216	140
217	140
218	141
219	141
220	142
221	142
222	142
223	143
224	143
225	144
226	144
227	145
228	145
229	145
230	146
231	146
232	147
233	147
234	148
235	148
236	148
237	149
238	149
239	150
240	150

KL-Werte	Standardwerte gesamt
241	151
242	151
243	151
244	152
245	152
246	153
247	153
248	154
249	154
250	154
251	155
252	155
253	156
254	156
255	157
256	157
257	157
258	158
259	158
260	159
261	159
262	160
263	160
264	160
265	161
266	161
267	162
268	162
269	163
270	163
271	163
272	164
273	164
274	165
275	165
276	166
277	166
278	166
279	167
280	167
281	168
282	168
283	169
284	169
285	169
286	170
287	170
288	171

Normstichprobe Schüler

$N = 312$

♂: $N = 185$, ♀: $N = 128$

≤ 20 Jahre: $N = 94$, 21 bis 25 Jahre: $N = 195$, 26 bis 30 Jahre: $N = 22$, ≥ 30 Jahre: $N = 2$

Fehler- %- Wert	Standardwerte		
	Form A	Form B	gesamt
0	82	83	82
1	85	87	85
2	89	90	89
3	92	93	82
4	95	97	96
5	99	100	99
6	102	104	103
7	105	107	106
8	108	110	110
9	112	114	113
10	115	117	117
11	118	121	120
12	121	124	124
13	125	127	127
14	128	131	131
15	131	134	134
16	134	138	138
17	138	141	141
18	141	144	145
19	144	148	148
20	148	151	152
21	151	155	155
22	154	158	159
23	157	161	162
24	161	165	166
25	164	168	169
26	167	172	173
27	170	175	176
28	174	178	180
29	177	182	183
30	180	185	187

Fehler- %- Wert	Standardwerte		
	Form A	Form B	gesamt
31	183	189	190
32	187	192	194
33	190	195	197
34	193	199	201
35	196	202	204
36	200	206	208
37	203	209	211
38	206	212	215
39	210	216	218
40	213	219	222
41	216	223	225
42	219	226	229
43	223	229	232
44	226	233	236
45	229	236	239
46	232	240	243
47	236	243	246
48	239	246	250
49	242	250	253
50	245	253	257
51	249	257	260
52	252	260	264
53	255	263	267
54	259	267	271
55	262	270	274
56	265	274	278
57	268	277	281
58	272	280	285
59	275	284	288
60	278	287	292
≥ 61	> 281	> 290	> 295

Normstichprobe Senioren

N = 136

figural ♂: N = 49, ♀: N = 87. numerisch und verbal ♂: N = 48, ♀ = 88
 ≤ 75 Jahre: N = 109, > 75 Jahre: N = 27

KL- Werte	Standardwerte					
	figural		numerisch		verbal	
	Form A	Form B	Form A	Form B	Form A	Form B
1	72	63	73	68	67	61
2	74	65	75	70	70	64
3	77	67	77	72	72	66
4	79	69	71	74	74	68
5	82	71	81	76	65	71
6	84	73	83	78	79	73
7	86	76	85	80	81	75
8	89	78	87	82	84	78
9	91	80	89	84	86	80
10	94	82	91	86	88	82
11	96	84	93	88	91	85
12	99	86	95	90	93	87
13	101	89	97	92	95	89
14	104	91	99	94	98	92
15	106	93	101	96	100	94
16	108	95	102	98	102	96
17	111	97	104	100	105	99
18	113	99	106	102	107	101
19	116	101	108	104	109	103
20	118	104	110	106	112	106
21	121	106	112	108	114	108
22	123	108	114	110	116	110
23	126	110	116	112	119	113
24	128	112	118	114	121	115
25	130	114	120	116	123	117
26	133	116	122	118	125	120
27	135	119	124	120	128	122
28	138	121	126	122	130	124
29	140	123	128	124	132	126
30	143	125	130	126	135	129
31	145	127	132	128	137	131
32	148	129	134	130	139	133
33	150	131	135	132	142	136
34	153	134	137	134	144	138
35	155	136	139	136	146	140
36	157	138	141	138	149	143
37	160	140	143	140	151	145
38	162	142	145	142	153	147
39	165	144	147	144	156	150
40	167	146	149	146	158	152
41	170	149	151	148	160	154
42	172	151	153	150	163	157
43	175	153	155	152	165	159
44	177	155	157	154	167	161
45	179	157	159	156	170	164
46	182	159	161	158	172	166
47	184	162	163	160	174	168
48	187	164	165	162	177	171

Normstichprobe Senioren

 $N = 136$

figural ♂: $N = 49$, ♀: $N = 87$. numerisch und verbal ♂: $N = 48$, ♀ = 88
 ≤ 75 Jahre: $N = 109$, > 75 Jahre: $N = 27$

KL- Werte	Standardwerte	
	Form A	Form B
1	62	56
2	63	57
3	64	58
4	65	58
5	66	59
6	67	60
7	68	61
8	69	62
9	70	63
10	71	64
11	71	64
12	72	65
13	73	66
14	74	67
15	75	68
16	76	69
17	77	70
18	78	70
19	79	71
20	80	72
21	81	73
22	81	74
23	82	75
24	83	75
25	84	76
26	85	77
27	86	78
28	87	79
29	88	80
30	89	81
31	90	81
32	91	82
33	91	83
34	92	84
35	93	85
36	94	86
37	95	86
38	96	87
39	97	88
40	98	89
41	99	90
42	100	91
43	101	92
44	101	92
45	102	93
46	103	94
47	104	95
48	105	96

KL- Werte	Standardwerte	
	Form A	Form B
49	106	97
50	107	97
51	108	98
52	109	99
53	110	100
54	111	101
55	112	102
56	112	103
57	113	103
58	114	104
59	115	105
60	116	106
61	117	107
62	118	108
63	119	109
64	120	109
65	121	110
66	122	111
67	122	112
68	123	113
69	124	114
70	125	114
71	126	115
72	127	116
73	128	117
74	129	118
75	130	119
76	131	120
77	132	120
78	132	121
79	133	122
80	134	123
81	135	124
82	136	125
83	137	125
84	138	126
85	139	127
86	140	128
87	141	129
88	142	130
89	142	131
90	143	131
91	144	132
92	145	133
93	146	134
94	147	135
95	148	136
96	149	136

KL- Werte	Standardwerte	
	Form A	Form B
97	150	137
98	151	138
99	152	139
100	152	140
101	153	141
102	154	142
103	155	142
104	156	143
105	157	144
106	158	145
107	159	146
108	160	147
109	161	148
110	162	148
111	162	149
112	163	150
113	164	151
114	165	152
115	166	153
116	167	153
117	168	154
118	169	155
119	170	156
120	171	157
121	172	158
122	172	159
123	173	159
124	174	160
125	175	161
126	176	162
127	177	163
128	178	164
129	179	164
130	180	165
131	181	166
132	182	167
133	183	168
134	183	169
135	184	170
136	185	170
137	186	171
138	187	172
139	188	173
140	189	174
141	190	175
142	191	175
143	192	176
144	193	177

Normstichprobe Senioren

$N = 136$

figural ♂: $N = 49$, ♀: $N = 87$. numerisch und verbal ♂: $N = 48$, ♀ = 88
 ≤ 75 Jahre: $N = 109$, > 75 Jahre: $N = 27$

KL-Werte	Standardwerte gesamt	KL-Werte	Standardwerte gesamt	KL-Werte	Standardwerte gesamt
1	57	49	79	97	101
2	58	50	80	98	101
3	58	51	80	99	102
4	59	52	80	100	102
5	59	53	81	101	103
6	60	54	81	102	103
7	60	55	82	103	103
8	61	56	82	104	104
9	61	57	83	105	104
10	61	58	83	106	105
11	62	59	84	107	105
12	62	60	84	108	106
13	63	61	84	109	106
14	63	62	85	110	107
15	64	63	85	111	107
16	64	64	86	112	108
17	65	65	86	113	108
18	65	66	87	114	108
19	66	67	87	115	109
20	66	68	88	116	109
21	66	69	88	117	110
22	67	70	89	118	110
23	67	71	89	119	111
24	68	72	89	120	111
25	68	73	90	121	112
26	69	74	90	122	112
27	69	75	91	123	113
28	70	76	91	124	113
29	70	77	92	125	113
30	70	78	92	126	114
31	71	79	93	127	114
32	71	80	93	128	115
33	72	81	94	129	115
34	72	82	94	130	116
35	73	83	94	131	116
36	73	84	95	132	117
37	74	85	95	133	117
38	74	86	96	134	117
39	75	87	96	135	118
40	75	88	97	136	118
41	75	89	97	137	119
42	76	90	98	138	119
43	76	91	98	139	120
44	77	92	99	140	120
45	77	93	99	141	121
46	78	94	99	142	121
47	78	95	100	143	122
48	79	96	100	144	122

Tabelle wird fortgesetzt

Normstichprobe Senioren

$N = 136$

figural ♂: $N = 49$, ♀: $N = 87$. numerisch und verbal ♂: $N = 48$, ♀ = 88

≤ 75 Jahre: $N = 109$, > 75 Jahre: $N = 27$

KL-Werte	Standardwerte gesamt
145	122
146	123
147	123
148	124
149	124
150	125
151	125
152	126
153	126
154	127
155	127
156	127
157	128
158	128
159	129
160	129
161	130
162	130
163	131
164	131
165	131
166	132
167	132
168	133
169	133
170	134
171	134
172	135
173	135
174	136
175	136
176	136
177	137
178	137
179	138
180	138
181	139
182	139
183	140
184	140
185	141
186	141
187	141
188	142
189	142
190	143
191	143
192	144

KL-Werte	Standardwerte gesamt
193	144
194	145
195	145
196	145
197	146
198	146
199	147
200	147
201	148
202	148
203	149
204	149
205	150
206	150
207	150
208	151
209	151
210	152
211	152
212	153
213	153
214	154
215	154
216	155
217	155
218	155
219	156
220	156
221	157
222	157
223	158
224	158
225	159
226	159
227	159
228	160
229	160
230	161
231	161
232	162
233	162
234	163
235	163
236	164
237	164
238	164
239	165
240	165

KL-Werte	Standardwerte gesamt
241	166
242	166
243	167
244	167
245	168
246	168
247	169
248	169
249	169
250	170
251	170
252	171
253	171
254	172
255	172
256	173
257	173
258	174
259	174
260	174
261	175
262	175
263	176
264	176
265	177
266	177
267	178
268	178
269	178
270	179
271	179
272	180
273	180
274	181
275	181
276	182
277	182
278	183
279	183
280	183
281	184
282	184
283	185
284	185
285	186
286	186
287	187
288	187

Normstichprobe Senioren

N = 135

♂: N = 48, ♀: N = 87

≤ 75 Jahre: N = 108, > 75 Jahre: N = 27

Fehler- %- Wert	Standardwerte		
	Form A	Form B	gesamt
0	83	86	83
1	85	88	86
2	88	91	89
3	90	93	92
4	93	96	94
5	95	99	97
6	98	101	100
7	101	104	102
8	103	106	105
9	106	109	108
10	108	111	110
11	111	114	113
12	113	117	116
13	116	119	118
14	118	122	121
15	121	124	124
16	123	127	127
17	126	130	129
18	128	132	132
19	131	135	135
20	133	137	137
21	136	140	140
22	138	143	143
23	141	145	145
24	143	148	148
25	146	150	151
26	148	153	154
27	151	156	156
28	153	158	159
29	156	161	162
30	158	163	164

Fehler- %- Wert	Standardwerte		
	Form A	Form B	gesamt
31	161	166	167
32	163	168	170
33	166	171	172
34	169	174	175
35	171	176	178
36	174	179	180
37	176	181	183
38	179	184	186
39	181	187	189
40	184	189	191
41	186	192	194
42	189	194	197
43	191	197	199
44	194	199	202
45	196	202	205
46	199	205	207
47	201	207	210
48	204	210	213
49	206	212	216
50	209	215	218
51	211	218	221
52	214	220	224
53	216	223	226
54	219	225	229
55	221	228	232
56	224	230	234
57	226	233	237
58	229	236	240
59	231	238	242
60	234	241	245
≥ 61	> 236	> 243	> 248

Lebenslauf

VERENA SCHMID

Persönliche Daten

Geburtsdaten 21. Juni 1969, Tulln, Österreich
Staatsangehörigkeit Österreich
Familienstand ledig

Schulbildung

1975 - 1979 Volksschule (Grundschule), Tulln, Österreich
1979 - 1983 Bundesgymnasium Tulln, Österreich
1983 - 1988 Handelsakademie Tulln, Österreich
Abschluss: Reifeprüfung, Note: 1.7

Studium

10/2002 - 09/2004 Ludwig-Maximilians-Universität, München
Grundstudium Psychologie, Diplom-Vorprüfung
Note: 1.56
10/2004 - 10/2007 Ludwig-Maximilians-Universität, München
Hauptstudium Psychologie, Schwerpunkte: Arbeits- und Organisationspsychologie (Projektseminar "Mitarbeiterbefragung") und Klinische Psychologie (Projektseminar "Diagnose und Therapie von Prüfungsängsten")
nicht-psychologisches Wahlpflichtfach: Psychopathologie
Abschluss: Diplom-Psychologin
Note: 1.13
04/2008 - 04/2010 Ludwig-Maximilians-Universität, München
Promotion im Hauptfach Psychologie unter der Leitung von Prof. Dr. Markus Bühner. Nebenfach: Psychopathologie
Titel der Dissertation: TEMEKKO - Validierung eines Tests zur Messung exekutiver Kontrolle / Konzentration.
Note: magna cum laude

beruflicher Werdegang

08/1988 - 07/1989	Samsung Electronics, Wien, Österreich Branch Secretary
08/1989 - 05/1991	Vienna Hilton, Wien, Österreich Direktionssekretärin
05/1991 - 08/1995	München City Hilton, München Direktionssekretärin
08/1995 - 10/1999	Time Life International GmbH, München Assistentin der Geschäftsführung
seit 11/1999	Assistentin von Prof. Dr. Alexander Hemmelrath, WP StB, München (bis 03/2007 Haarmann Hemmelrath, seit 04/2007 Mazars Hemmelrath)

Praktika

07/1986	Creditanstalt-Bankverein, Wien Volontariat, Akkreditiv-Abteilung
07/1987	Creditanstalt-Bankverein, Wien Volontariat, Akkreditiv-Abteilung
12/2005 - 06/2006	Haarmann Hemmelrath, München Human Resources/Personalentwicklung, Change Management

berufliche Qualifikationen

06/1987	WIFI Niederösterreich Steno- und Phonotypieprüfung
07/1999	Bernecker & Partner, Trainingsseminar 560 NetWare 5.x Administration
03/2005	Ludwig-Maximilians-Universität, Prof. Dr. Dieter Frey Seminar "Mitarbeiterorientierte Unternehmensführung"

Vortrag

09/2009

Validierung eines Markertests zur Messung exekutiver Kontrolle/Konzentration. 10. Arbeitstagung der Fachgruppe Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik der Deutschen Gesellschaft für Psychologie, Landau, 27.-30. September 2009 (Schmid, V., Bühner, M., Schmidt-Atzert, L. & Ziegler, M.)
