



universität  
wien

## Diplomarbeit

Titel der Diplomarbeit

**Strukturierungs- und Gliederungsfähigkeit im Untertest  
„Analysieren und Synthetisieren“ des AID 3:  
Identifikation der kognitiven Operationen mittels LLTM**

Verfasserin

Sabrina Köb

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2013

Studienkennzahl: A 298

Studienrichtung: Psychologie

Betreuer: Univ.-Prof. Mag. Dr. Klaus D. Kubinger



# Danksagung

---

Die Arbeit, die in den folgenden Seiten steckt, hat sich über einen langen Zeitraum hingezogen, in dem viele Nerven verloren und neue Erkenntnisse gewonnen wurden. Die Diplomarbeit hat sich aber nicht nur als das größte und aufwändigste, sondern glücklicher Weise auch als das interessanteste Projekt meines Studiums herausgestellt.

Für die Möglichkeit, mich in ein Thema zu vertiefen, das mich seit Beginn des Studiums fasziniert hat, sowie für das große Vertrauen in mein eigenständiges Arbeiten, möchte ich mich bei meinem Betreuer, Univ. Prof. Dr. Mag. Klaus D. Kubinger, bedanken. Ich konnte im letzten Jahr durch seinen großen Erfahrungs- und Wissensschatz noch viel dazu lernen.

Nichtsdestotrotz war aber auch ein großer eigenständiger Lernaufwand vonnöten. Wenn trotz hohem Einsatz einmal die Motivation zu schwinden drohte, fand ich oft in den Worten und Taten meiner Nächsten den nötigen kleinen oder großen Anstoß, um effizient weiter machen zu können.

Deshalb möchte ich mich bei meinen Studienkolleginnen und -kollegen für den regen Austausch und die große Hilfsbereitschaft bedanken, sowie bei meinen Freunden und Freundinnen in- und außerhalb des Studiums für die offenen Ohren und die motivierenden Worte. Einige Personen möchte ich diesbezüglich speziell erwähnen.

Ronja,

... danke, nicht nur für deinen laufenden Beistand und die wertvolle Hilfe beim Korrekturlesen, sondern auch für die aufmunternden Pausengespräche bei langen Tagen in der Bibliothek.

Claudia,

... danke für deine Unterstützung beim Korrektur lesen, vor allem aber für deine fürsorgliche Art, mich bei kleinen wie großen Krisen zu beruhigen und wieder aufzubauen.

Armin,

... danke für den 24/7-Technik-Notfall-Dienst, den ich mehr als nur ein Mal in Anspruch genommen habe.

Ganz besonders herzlicher Dank gilt meinen Eltern, die sich in ihrer Bereitschaft, mich zu unterstützen, und besonders in Sachen Geduld, selbst übertroffen haben.

Ihr habt mir stets alle Türen offen gehalten und mir dieses Studium ermöglicht. Ihr wart und seid mir in allen Lebenslagen die wichtigste Stütze.  
Euch ist diese Arbeit gewidmet.

*Ohne eu wär's nid ganga.*



## Zusammenfassung

Ziel der vorliegenden Arbeit war es erstens, jene mentalen Operationen zu identifizieren, welche für die Lösung der Aufgaben im Untertest *Analysieren und Synthetisieren* des AID 3 (Kubinger & Holocher-Ertl, in Druck) nötig sind, und zweitens, die entsprechenden Annahmen mittels *Linear Logistischem Testmodell (LLTM)* (Fischer, 1973) empirisch zu prüfen.

Basierend auf theoretischen Annahmen, praktischen Erfahrungen mit dem Test sowie einer sorgfältigen Itemanalyse wurden Hypothesen über den Prozess des mentalen Strukturierens in Form von Elementaroperationen in der Strukturmatrix des *LLTM* definiert.

Als Berechnungsgrundlage dienen Normierungstaten für den AID 3 ( $n = 1130$ ). Die Geltung des *Rasch*-Modells (Rasch, 1980) als Voraussetzung für das *LLTM* kann bestätigt werden. Eine erste Berechnung des Modells führt zu enttäuschenden Ergebnissen des *Likelihood Ratio Tests (LRT)* (Andersen, 1973;  $\chi^2_{emp} = 595,971$ ,  $\chi^2_{krit(0,05;10)} = 18,307$ ). Durch die Reduktion des Itempools um ein leicht auffälliges Item wird dessen Homogenität erhöht und die Abweichung der Itemparameterschätzungen des *LLTM* von jenen des *Rasch*-Modells reduziert. Die graphische Modellkontrolle zeigt eine hohe Übereinstimmung für 16 der 17 Schätzwerte. Der *LRT* führt jedoch abermals zu einer Verwerfung des Modells ( $\chi^2_{emp} = 251,924$ ,  $\chi^2_{krit(0,05;10)} = 16,919$ ).

Mögliche Gründe für die mangelnde Vorhersagekraft des Modells werden diskutiert und Verbesserungsvorschläge geboten. Trotz der kritischen Betrachtung des entworfenen Modells im Ganzen wird ein Nutzen als Grundlage für weitere Untersuchungen aufgezeigt. Bei der Betrachtung der geschätzten Basisparameter scheinen die Hypothesen über die definierten Elementaroperationen zumindest teilweise bestätigt.

## Abstract

The goal of this study was to identify the cognitive operations necessary for solving the items of the AID 3 subtest *Analysieren und Synthetisieren* (Kubinger & Holocher-Ertl, in print) via the use of the *LLTM* (*Linear Logistic Test Model*; Fischer, 1973).

Based on theoretical assumptions, practical experiences with the test and a thorough item analysis, assumptions about the use of a mental grid and rotation for solving the tasks were defined as elementary operations in the *LLTM* structure matrix.

With data from the standardization sample of the AID 3 ( $n = 1130$ ) a *Rasch Model (RM)* (Rasch, 1980) fitting item pool could be confirmed. First calibrations of the *LLTM* led to disappointing results of the *Likelihood Ratio Test (LRT)* (Andersen, 1973;  $\chi^2_{emp} = 595,971$ ,  $\chi^2_{krit(0,05;10)} = 18,307$ ). In eliminating a rather critical item from the item pool, not only a better fit of the *RM* in general, but also a better fit of the *LLTM* item parameter estimates to the *RM* estimates is established. A graphical model check shows that the parameter estimates for 16 of the 17 items are quite accurate. Only item 6 is evidently not to be explained by the model. Still the null hypothesis of the *LLTM* has to be discarded due to the results of the *LRT* which are again significant ( $\chi^2_{emp} = 251,924$ ,  $\chi^2_{krit(0,05;10)} = 16,919$ ).

Possible reasons for this result are discussed and suggestions for improvement are offered. Despite the critical review of the model its usefulness for further studies is assumed. Also, in view of the basic parameter estimates, which correspond in general with the hypotheses of this study, the relevance of certain here established elementary operations for explaining item difficulty in the AID subtest *Analysieren und Synthetisieren* seems to be confirmed.



# Inhaltsverzeichnis

---

<b>I. Einleitung.....</b>	<b>1</b>
<b>II. Untertest <i>Analysieren und Synthesieren-abstrakt (UT 10)</i> im AID 3 ....</b>	<b>3</b>
1. Aufgabenstellung und Material des UT 10.....	4
2. Validität des UT 10.....	5
3. Mit dem UT 10 verwandte Verfahren.....	5
3.1 Der <i>Blockdesign</i> -Test .....	6
3.2 Untertest <i>Strukturieren (UT 10a)</i> im AID 3 .....	7
3.3 Der Gestaltwahrnehmungstest .....	7
<b>III. Theoretischer Rahmen der Strukturierungsfähigkeit .....</b>	<b>9</b>
4. Experimentelle Studien mit verwandten Verfahren.....	9
4.1 Lösungsprozesse und Mustereigenschaften.....	10
4.2 Bearbeitungsstrategien .....	13
5. Zur kognitiven Organisation visueller Wahrnehmungsinformationen .....	15
5.1 Die Gestaltpsychologie und ihre Gesetze.....	17
5.1.1 Prinzipien der Figur-Grund-Trennung .....	18
5.1.2 Wahrnehmungsorganisation im Kindesalter .....	20
5.1.3 Die gute Gestalt .....	20
5.1.4 Operationalisierung von guter Gestalt mittels <i>SIT</i> .....	22
5.2 Das Konzept Feldabhängigkeit/Feldunabhängigkeit .....	23
5.2.1 Historischer Überblick von Feldabhängigkeit/ Feldunabhängigkeit .....	24
5.2.2 Entwicklungsaspekte der Feldabhängigkeit/Feldunabhängigkeit.....	26
5.2.3 Zur Dimensionalität von Feldabhängigkeit/Feldunabhängigkeit .....	27

5.3	Freiheit von Gestaltbindung.....	28
6.	Schlussfolgernde Zusammenfassung.....	30
<b>IV. Testtheoretischer Rahmen des linear logistischen Testmodells .....</b>		<b>32</b>
7.	Das dichotom logistische Testmodell von Rasch ( <i>RM</i> ).....	32
8.	Die Prüfung der Geltung des <i>RM</i> .....	34
8.1	Der <i>Likelihood Ratio Test</i> nach Andersen beim <i>RM</i> .....	34
8.2	Die graphische Modellkontrolle beim <i>RM</i> .....	35
9.	Das linear logistische Testmodell ( <i>LLTM</i> ) von Fischer .....	36
9.1	Die Strukturmatrix des <i>LLTM</i> ( <i>Q</i> -Matrix).....	38
9.2	Zur Anwendung des <i>LLTM</i> .....	39
10.	Die Prüfung der Geltung des <i>LLTM</i> .....	41
10.1	Der <i>Likelihood Ratio Test</i> nach Andersen beim <i>LLTM</i> .....	41
10.2	Die graphische Modellkontrolle beim <i>LLTM</i> .....	42
<b>V. Modellderstellung.....</b>		<b>43</b>
11.	Vorgehensweise.....	43
12.	Modellrahmen und Hypothesengenerierung .....	44
13.	Hypothesen und Elementarfaktoren .....	46
14.	Das Modell in Form der <i>Q</i> -Matrix .....	53
<b>VI. Modellprüfung.....</b>		<b>54</b>
15.	Stichprobe und Datensatz.....	54
16.	Qualität der <i>Q</i> -Matrix .....	54
17.	Modellprüfungen .....	55
17.1	Prüfung der <i>Rasch</i> -Modell-Konformität aller Items .....	55
17.2	Prüfung der Geltung des <i>LLTM</i> mit allen Items .....	57

17.3	Prüfung der <i>Rasch</i> -Modell-Konformität nach Ausschluss von Item 18 .....	59
17.4	Prüfung der Geltung des <i>LLTM</i> nach Ausschluss von Item 18 .....	59
<b>VII.</b>	<b>Diskussion .....</b>	<b>62</b>
<b>VIII.</b>	<b>Zusammenfassung.....</b>	<b>66</b>
<b>IX.</b>	<b>Literaturverzeichnis .....</b>	<b>69</b>
<b>X.</b>	<b>Anhang.....</b>	<b>76</b>
1.	Testperson-Notizen ( <i>TPN</i> ) .....	76
2.	Testleiter-Notizen ( <i>TLN</i> ).....	79
3.	Lebenslauf.....	83

# **Abbildungsverzeichnis**

---

Abbildung 1: Oberflächenseiten der Würfel im Untertest 10 <i>Analysieren und Synthetisieren</i> ..	4
Abbildung 2: Beispiel eines möglichen Streifenmusters mit diagonal geteilten Würfelseiten	13
Abbildung 3: Beispiel eines möglichen Rautenmusters mit diagonal geteilten Würfelseiten .	13
Abbildung 4: Beispiel für die graphische Kontrolle des Rasch-Modells.....	36
Abbildung 5: graphische Kontrolle des <i>RM</i> mit allen Items unter Einbezug des Konfidenzbandes ( $\alpha = 0,01$ ) für das Trennkriterium Geschlecht .....	56
Abbildung 6: graphische Kontrolle des <i>RM</i> mit allen Items unter Einbezug des Konfidenzbandes ( $\alpha = 0,01$ ) für das Trennkriterium Staat .....	57
Abbildung 7: graphische Kontrolle des <i>LLTM</i> mit allen 18 Items .....	58
Abbildung 8: graphische Kontrolle des <i>LLTM</i> nach Ausschluss von Item 18.....	60

# Tabellenverzeichnis

---

Tabelle 1: Schematische Darstellung der Gewichtszahlenmatrix des <i>LLTM</i> .....	39
Tabelle 2: Definition und Gewichtung von Elementarfaktor <i>Ecke</i> .....	48
Tabelle 3: Definition und Gewichtung von Elementarfaktor <i>Kanten</i> .....	49
Tabelle 4: Definition und Gewichtung von Elementarfaktor <i>Kreuzpunkte</i> .....	50
Tabelle 5: Definition und Gewichtung der Elementaroperation <i>dritteln</i> .....	51
Tabelle 6: Definition und Gewichtung von Elementarfaktor <i>Kreuzpunkte</i> .....	51
Tabelle 7: Definition und Gewichtung von Elementarfaktor <i>Symmetrie</i> .....	52
Tabelle 8: Definition und Gewichtung der Elementaroperation <i>rotieren</i> .....	53
Tabelle 9: vollständige Q-Matrix mit den sieben Elementarfaktoren <i>Ecke, Kanten,</i> <i>Kreuzpunkte, dritteln, , Hintergrund, Symmetrie und rotieren</i> .....	53
Tabelle 10: Korrelationsmatrix mit den sieben Elementarfaktoren <i>Ecke, Kanten, Kreuzpunkte,</i> <i>dritteln, , Hintergrund, Symmetrie und rotieren</i> .....	55
Tabelle 11: Ergebnisse der RM-Prüfung anhand des LRT nach Andersen für die Trennkriterien Geschlecht und Staat .....	56
Tabelle 12: Ergebnis der <i>LLTM</i> -Prüfung mittels <i>LRT</i> nach Andersen .....	57
Tabelle 13: Schätzungen der Basisparameter mit Schätzfehler ( <i>SE</i> ) im <i>LLTM</i> mit allen Items .....	58
Tabelle 14: Ergebnisse der RM-Prüfung nach Ausschluss von Item 18 anhand des <i>LRT</i> nach Andersen für die Trennkriterien Geschlecht und Staat.....	59
Tabelle 15: Ergebnisse der <i>LLTM</i> -Prüfungen mittels <i>LRT</i> nach Andersen vor und nach Ausschluss von Item 18 .....	60
Tabelle 16: Schätzungen der Basisparameter Schätzfehler ( <i>SE</i> ) im <i>LLTM</i> nach Ausschluss von Item 18 .....	61
Tabelle 17: Gegenüberstellung der Itemparameter ( <i>σi</i> ) des <i>RM</i> und des <i>LLTM</i> nach Ausschluss von Item 18 mit dazugehörigen Schätzfehlern ( <i>SE</i> ).....	61
Tabelle 18: Testperson-Notizen ( <i>TPN</i> ) über introspektive Beobachtungen, auffällige Merkmale der Materialeigenschaften sowie darauf begründete Annahmen zum Lösungsprozess bei der Bearbeitung des UT 10 in der Position als Testperson... 76	76
Tabelle 19: Testleiter-Notizen ( <i>TLN</i> ) über Gelegenheitsbeobachtungen und darauf begründete Annahmen zum Lösungsprozess bei der Vorgabe der Untertests 10 und 10a..... 79	79



# I. Einleitung

---

Neben Objektivität und Reliabilität gehört die Validität zu den Hauptgütekriterien, welche für psychologisch-diagnostische Verfahren gelten sollen (Lienert & Raatz, 1998). Valide ist ein psychologischer Test dann, wenn er „tatsächlich jenes psychische Merkmal misst, welches er zu messen behauptet“ (Kubinger, 2009b, S. 55). Erfordert die Bearbeitung einer Aufgabe die Anwendung bestimmter kognitiver Operationen und ist die Lösung auf keine andere Weise zu erzielen, so kann von einem hohen Testergebnis auf eine hohe Kompetenz in der Durchführung der entsprechenden kognitiven Operationen geschlossen werden (Kubinger, 2008).

Ziel der vorliegenden Arbeit ist es, für die zweidimensionalen Items des Untertests *Analysieren und Synthetisieren* (UT 10) im *Adaptiven Intelligenz Diagnostikum 3* (AID 3; Kubinger & Holocher-Ertl, in Druck) die zur Lösung erforderlichen kognitiven Operationen zu identifizieren und bezüglich deren Beitrag zur Itemschwierigkeit zu überprüfen. Das Konstrukt, für welches der Test erstellt wurde, ist die Fähigkeit, „komplexe (abstrakte) Gestalten durch eine geeignete Strukturierung reproduzieren zu können“ (Kubinger, 2009a, S. 11). Diese erste Definition des mentalen Lösungsprozesses bildet den Ausgangspunkt für eine weitere Aufschlüsselung in einzelne mentale Operationen. Die Herausforderung liegt dabei im Identifizieren und Definieren eindeutig abgrenzbarer kognitiver Operationen und deren Beziehung zu bestimmten Eigenschaften der Items. Darauf aufbauend soll eine Strukturmatrix erstellt werden, welche Informationen über das Vorliegen von Anforderungen an bestimmte kognitive Operationen je Item enthält. Dies bildet die Grundlage für die Berechnung des *Linear Logistischen Testmodells* (Fischer, 1973). Bei dessen Gültigkeit können die Itemschwierigkeiten durch die Notwendigkeit bestimmter kognitiver Operationen genau vorhergesagt werden.

Den Kern der Hypothesengenerierung bilden

- a) die sorgfältige Betrachtung des Materials,
- b) die Erkenntnisse durch Studien zu ähnlichen Verfahren, sowie
- c) die intensive Auseinandersetzung mit dem theoretischen Rahmen zur Strukturierungs- und Gliederungsfähigkeit anhand einschlägiger Literatur.

Für letzteres werden einerseits in Hinblick auf einen möglichen Beitrag zur Erfassung von schwierigkeitssteigernden und schwierigkeitssenkenden Materialeigenschaften ausgewählte Prinzipien der Gestaltpsychologie besprochen und andererseits, als dem UT 10 womöglich zugrundeliegende Konstrukte, das Konzept der Feldabhängigkeit/Feldunabhängigkeit sowie der Faktor *Freiheit von Gestaltbindung* in Betracht gezogen und diskutiert.

Vor Beginn der eigentlichen Arbeit wurde das Material des UT 10 noch unvoreingenommen von konkretem Vorwissen über die anzunehmenden kognitiven Prozesse und relevanten Itemcharakteristika betrachtet. Aus Erfahrungen als Testperson sowie Gelegenheitsbeobachtungen als Testleiter konnte eine Sammlung an Notizen erstellt werden. Obwohl diese nicht durch eine systematische Erhebung und Auswertung entstanden sind, sollen sie der Vollständigkeit halber angeführt und für das Aufstellen der Hypothesen ergänzend berücksichtigt werden.

Der obligatorischen Überprüfung der *Rasch*-Modell-Konformität der Daten folgt die Berechnung des Linear Logistischen Testmodells anhand der entwickelten Strukturmatrix. Bei Gültigkeit kann davon ausgegangen werden, dass durch die mentalen Operationen die Schwierigkeit der Items angemessen genau vorhergesagt werden können.

#### **Anmerkung:**

Im Sinne einer leichteren Lesbarkeit wird in dieser Arbeit auf eine geschlechtsspezifische Formulierung verzichtet. Auf Personen bezogene Bezeichnungen stehen, soweit es sinnvoll ist, für Frauen und Männer in gleicher Weise.

## **II. Untertest *Analysieren und Synthetisieren*-abstrakt (UT 10) im AID 3**

---

Das *Adaptive Intelligenz Diagnostikum 3* (AID 3) von Kubinger und Holocher-Ertl (in Druck) ist die jüngste Version der förderdiagnostisch orientierten Intelligenz-Testbatterie für Kinder und Jugendliche im Alter von 6;0 bis 15;11 Jahren, welche erstmals im Jahre 1985 (Kubinger & Wurst) vorgestellt wurde. Sie setzt sich aus zwölf Untertests sowie drei Zusatztests zusammen. Grundlage für die Konzipierung der einzelnen Sub- und Zusatztests waren operationale Definitionen. Es wurde also begrifflich exakt festgelegt, was zu erfassen beabsichtigt ist.

„Der Untertest **10 Analysieren und Synthetisieren**-abstrakt soll die Fähigkeit prüfen, komplexe (abstrakte) Gestalten durch eine geeignete Strukturierung reproduzieren zu können“ (Kubinger, 2009a, S. 11). Annahmen über diese Strukturierungsleistung und die dafür relevanten Itemeigenschaften werden daher den Kern der Arbeit bilden.

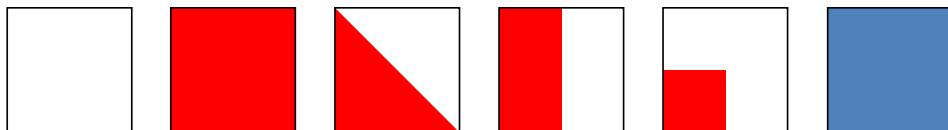
Der Untertest wird im Einzelsetting *adaptiv* im Sinne des *branched testing* vorgegeben. Voraussetzung für das adaptive Testen ist die Gültigkeit des sogenannten *Rasch*-Modells, welches in Abschnitt 7 genauer besprochen wird. Aufgaben, die einer Testperson sehr leicht oder sehr schwer fallen, sind für die Einschätzung ihrer Fähigkeit wenig informativ, da das Ergebnis der Bearbeitung von vornherein recht klar ist. Beim *Rasch*-Modell sind die jeweiligen Schwierigkeiten der Items bekannt. Das erlaubt es, durch deren Bearbeitung Rückschlüsse auf die Fähigkeit der Testperson zu ziehen und bereits nach wenigen Items eine erste grobe Einstufung der Fähigkeit zu treffen. Indem die Schwierigkeiten der folgenden Items entsprechend dieser Einschätzung gewählt werden, bearbeitet die Testperson keine Aufgaben, die ihr sehr leicht oder sehr schwer fallen. Beim *branched testing* wird hierfür ein bestimmtes Verzweigungsschema angewandt. So kann auf sehr ökonomische Weise eine genaue Schätzung erzielt werden (Kubinger & Holocher-Ertl, in Druck).

Der Untertest *10 Analysieren und Synthetisieren* (UT 10) wird den manuell-visuellen Fähigkeiten zugeordnet und besteht aus 18 zweidimensionalen sowie sieben dreidimensionalen Items. Durch die Vorgabe entsprechend des *branched testing* hat jedes

Kind 6 Aufgaben zu bearbeiten. Die drei Startgruppen sind altersspezifisch für die 6- bis 9-Jährigen, die 10- und 11-Jährigen sowie die 12- bis 15-Jährigen aufgeteilt. Ein abgeändertes Verzweigungsschema sowie die dreidimensionalen Items sind Neuerungen gegenüber seiner Vorgängerversion im AID 2, welche nur aus den 18 zweidimensionalen Aufgaben bestand (Kubinger & Holocher-Ertl, in Druck). Da für die vorliegende Arbeit lediglich die zweidimensionalen Items von Bedeutung sind, sollen die dreidimensionalen Items in der weiteren Beschreibung nicht speziell berücksichtigt werden.

## 1. Aufgabenstellung und Material des UT 10

Den Kindern und Jugendlichen wird auf einer Karte eine Vorlage präsentiert, welche je ein weiß-rotes Muster abbildet. Aufgabe ist es, dieses Muster anhand von Würfeln mit verschiedenen Musterelementen je Würfelseite nachzulegen. Die neun Würfel sind ident und haben je fünf für die Lösung relevante Seiten, wobei drei davon zweifarbig sind, sowie eine blaue, für die Lösung irrelevante, Seite. (Abbildung 1). Die nachzulegenden Muster setzen sich aus zwei, vier oder neun Würfelseiten zusammen. Den Kindern wird die entsprechende Anzahl an Würfeln je Aufgabe zufällig verteilt hingelegt.



**Abbildung 1:** Oberflächenseiten der Würfel im Untertest 10 *Analysieren und Synthetisieren*  
Jeder der neun Würfel hat dieselben Seiten. Die Würfelseiten werden von links nach rechts folgendermaßen bezeichnet: *weiß* (*w*), *rot* (*r*), *diagonal geteilt* (*d*), *horizontal/vertikal geteilt* (*hv*), *viertel Ecke* (*v*), *blau*

Dem eigentlichen Testteil geht die Vorgabe von 2 *warming-up* Aufgaben voran. Die Lösung der Aufgaben muss innerhalb bestimmter Zeitbeschränkungen geschehen, welche je nach Schwierigkeit des Items zwischen 30 und 90 Sekunden liegen. Die Wertung für ein zeitgerecht nachgelegtes Muster ist „1“, ansonsten wird das Item mit „0“ als *nicht gelöst* verrechnet.

## **2. Validität des UT 10**

Für die differenzialdiagnostische Einschätzung von Teilleistungsstörungen ist der UT 10 besonders im Bereich der Differenzierungs- und Gliederungsfähigkeit von Nutzen. Dies betrifft das differenzierte Wahrnehmen ähnlicher Reizinhalte sowie die Gliederung und Trennung komprimierter Informationen. Aber auch auf die Raum-Lage-Orientierung, als Form der visuellen Gliederungsfähigkeit, sowie auf das Verarbeiten und das benutzen einmal eingeprägter Inhalte können aus den Ergebnissen Rückschlüsse gezogen werden (Kubinger & Holocher-Ertl, in Druck).

Durch die spezielle Konstruktion der Untertests auf Basis operanter Definitionen gelingt auf herkömmliche Weise lediglich die diskriminante Konstruktvalidierung. Die im Manual des AID 2 berichteten Untersuchungsergebnisse besagen, dass der UT 10 nicht Raumvorstellung erfasst, wie beispielsweise der 3DW (*Dreidimensionaler Würfeltest*; Gittler, 1990), wenngleich von einer gewissen Anforderung an die Raum-Lage-Orientierung ausgegangen werden kann. Auch Reasoning im nonverbalen, anschauungsgebundenen Bereich wird nicht direkt erfasst, wenngleich von einigen niedrigen bis mittleren Zusammenhängen berichtet wird (AID 2).

## **3. Mit dem UT 10 verwandte Verfahren**

Kohs' (1920) *Blockdesign*-Test kann im weiteren Sinne als Vorgänger des UT 10 gesehen werden. Das als Intelligenztest konzipierte Verfahren erfreute sich großer Beliebtheit und wurde in abgeänderter Form in mehrere Intelligenz-Testbatterien integriert – so zum Beispiel neben dem AID auch in den Wechlser-Intelligenz-Testbatterien WISC (Wechlser Intelligence Scales for Children; Wechsler, 2003) und WAIS (Wechlser Adult Intelligence Scales; Petermann, 2012).

Der Zusatztest *10a Strukturieren* im AID 3 weist, wie bereits die Nummerierung vermuten lässt, viele Gemeinsamkeiten mit dem UT 10 auf und soll deshalb ebenfalls einer genaueren Betrachtung unterzogen werden. Weniger durch Aufgabenstellung und Material, als durch ein gemeinsames zugrundeliegendes latentes Konstrukt, erwies sich der *Gestaltwahrnehmungstest* als dem UT 10 verwandt (Hergovich, 1992). In diesem Verfahren

müssen die Testpersonen eine bestimmte Form, das Häuschen, in einem komplexen Liniengefüge wiederentdecken.

### 3.1 Der *Blockdesign*-Test

Der *Blockdesign*-Test ist bezüglich des Materials, der Aufgabenstellung und der Vorgabe eng mit dem UT 10 verwandt. Er wurde 1920 erstmals von Kohs als sprachfreier Intelligenztest publiziert. Die Vorgabemuster (*designs*) weisen eine sehr ähnliche Struktur auf (tatsächlich ist das Muster 1.3 des UT 10 ident mit dem Design 5 des *Blockdesign* und Design 4 unterscheidet sich von 5.1 nur durch die Farbkombination). Die Würfelseiten (*blocks*) haben im Vergleich zum UT 10 mehrere Farben doch die Musterelemente beschränken sich auf volle und diagonal geteilte Würfelseiten. Jeder der insgesamt 16 Würfel hat je eine blaue, eine rote, eine weiße und eine gelbe Seite sowie eine diagonal geteilte rot-weiße und eine diagonal geteilte blau-gelbe Seite. Die Muster sind zentriert im Maßstab 1:4 auf dem Vorlagekärtchen gedruckt. Die Beobachtung und Verrechnung verläuft beim *Blockdesign* allerdings weitaus anders als beim UT 10, denn sowohl die Zeit bis zur Lösung als auch die Anzahl an Zügen (*moves*) werden mit in den Testwert einbezogen. Damit sollen zusätzlich Geschwindigkeit und Genauigkeit miterfasst werden.

#### Block Design im WISC

Ähnlich wie im AID wurde bereits in die Wechsler Intelligence Scales for Children (WISC), für die mittlerweile die vierte Version (WISC-IV; Wechsler, 2003) vorliegt, die *Blockdesign*-Aufgaben in Form des Subtests *Block Design* integriert. Dieser wird in Wechslers Konzept dem wahrnehmungsgebundenen logischen Denken zugeordnet. Während die WISC-IV für Kinder und Jugendliche von 6 bis 16 Jahren gedacht ist, stellt die Wechsler Adult Intelligence Scales (WAIS; Petermann, 2012 für die aktuelle deutsche Version WAIS-IV) das Pendant für Jugendliche ab 16 Jahren und Erwachsene dar.

Anders als der originale *Blockdesign*-Test oder der UT 10 im AID besteht dieser Untertest jedoch lediglich aus roten, weißen und diagonal geteilten rot-weißen Würfelseiten. Jeder Würfel hat daher je zwei solche Seiten. Die nachzulegenden Muster bestehen entweder aus vier oder neun Würfeln.

### **3.2 Untertest Strukturieren (UT 10a) im AID 3**

„Der Zusatztest **10a Strukturieren**-visumotorisch soll die Fähigkeit erfassen, komplexe (abstrakte) Gestalten in elementare Teilkomponenten zerlegen zu können.“ (Kubinger, 2009a, S. 13)

Der Zusatztest ist für Kinder ab 8 Jahren gedacht. Es werden Muster vorgelegt, welche sich aus denselben Würfelseiten wie die Muster im UT 10 zusammensetzen lassen. Allerdings sind die Muster hier nicht auf eine quadratische Zusammensetzung beschränkt und bestehen aus unterschiedlichen Anzahlen an Würfelseiten.

Aufgabe ist es, mit einem Stift auf der Vorlage jene Linien einzuziehen, welche sich ergeben, wenn man das Muster mit den Würfeln nachlegen würde. Es gilt also die Kanten der Würfel bzw. die Ränder der Würfelseiten in die jeweiligen Muster einzufügen (Kubinger, 2009a).

Insgesamt umfasst der Zusatztest eine Beispielaufgabe und 11 Items. Die Vorgabe erfolgt konventionell und ist auf 2 Minuten beschränkt. Es werden zeitgleich alle 11 Muster auf dem Testbogen präsentiert. Die Reihenfolge deren Bearbeitung wird dabei den Kindern überlassen. Wie der Untertest 10 wird auch Zusatztest 10a für die Einschätzung der Differenzierungs- und Gliederungsfähigkeit herangezogen, wobei letzterer weniger Ansprüche an die Motorik stellt. Hierdurch sollen Störungen der Wahrnehmung von Störungen der Verarbeitung auseinander gehalten werden können (Kubinger, 2009a).

### **3.3 Der Gestaltwahrnehmungstest**

1965 wurde von Vukovich und Mitarbeitern der *Häuschentest* entwickelt (unveröffentlicht, zitiert nach Dieterich, 1975). Hergovich (1992) entwickelte darauf aufbauend ein computerisiertes Verfahren, welches als *Gestaltwahrnehmungstest* (GESTA) veröffentlicht wurde (Hergovich & Hörndl, 1994). Es enthält mehrere komplexe Figuren, in denen jeweils der Umriss eines schematisierten Häuschens enthalten ist, welcher gefunden und nachgezeichnet werden muss. Die Aufgaben sind verschieden schwierig und müssen innerhalb einer bestimmten Zeit gelöst werden. Die Suchfigur, also das Häuschen, bleibt bei der Bearbeitung stets gleichzeitig dargeboten.

Es wird zwischen strukturierten und unstrukturierten Figuren unterschieden. Die strukturierten Figuren sind regelmäßiger und häufig symmetrisch. Ihnen liegt eine erkennbare Systematik zugrunde. Sie sind in diesem Sinne bessere Gestalten (vgl. Wertheimer, 1923), da sie als prägnante Einheit bzw. Figur wahrgenommen werden. Die unstrukturierten Figuren hingegen wirken eher unausgewogen und zufällig (Dieterich, 1975).

Der Test gilt als Verfahren zur Erfassung von Feldabhängigkeit/Feldunabhängigkeit. Sowohl für die Papier-Bleistift-Version des *Häuschentests* als auch für die Computer-Version konnte Eindimensionalität nachgewiesen werden (Hergovich, 1992; Waldenmair, 1984). Von besonderer Bedeutung für die vorliegende Arbeit sind die Ergebnisse von Hergovich (1992), wonach einige Items des UT 10 aus der ersten Version des AID mit den Aufgaben des *Häuschentests* eindimensional messen.

## **III. Theoretischer Rahmen der Strukturierungsfähigkeit**

---

Für die theoretische Fundierung wird sich einem Ausschnitt des breiten Themenfelds zur Wahrnehmungsorganisation gewidmet. Die Prinzipien der Gestalttheorie sollen in Bezug auf deren potentiellen Beitrag für die adäquate Erfassung der Materialeigenschaften des UT 10 beleuchtet werden. Das Konzept der Feldabhängigkeit/Feldunabhängigkeit und der Faktor *Freiheit von Gestaltbindung* sollen in Hinblick auf deren Übereinstimmung einerseits mit der operanten Definition des UT 10 sowie andererseits mit gewissen Aspekten der dazugehörigen Verfahren einer genaueren Betrachtung unterzogen werden.

Bevor sich jedoch diesen vergleichsweise allgemein formulierten Themen gewidmet wird, soll ein Überblick über empirische Arbeiten gegeben werden, die sich sehr spezifisch mit den Lösungsprozessen für *Blockdesign*-Aufgaben befassen.

### **4. Experimentelle Studien mit verwandten Verfahren**

Kohs (1920) zitiert für die Bestimmung der kognitiven Beanspruchung beim Lösungsprozess im *Blockdesign*-Test Binet (1916, zitiert nach Kohs, 1920) wonach das Zieldesign verstanden und im Sinn behalten werden muss, verschiedene Kombinationen entsprechend einer bestimmten Idee ausprobiert werden und es gilt, das Ergebnis zu beurteilen, zu vergleichen und als passend zu identifizieren. Kohs argumentiert des Weiteren, dass – ebenfalls entsprechend Binet – Aufmerksamkeit, „*adaptation*“ und Selbstkritik zur Lösung vonnöten sind. Zur Definition des *Blockdesign*-Test als Intelligenztest listet Kohs zudem folgende mentale Operationen:

If 'intelligence' involves the following mental operations; analyzing, combining, comparing, deliberating, completing, discriminating, judging, criticising and deciding, then the block-design tests may, with justice, be said to call upon the functioning of intelligence and to that extent they are a measure of that mental capacity. (S. 370)

Die Items im *Blockdesign* wurden von Kohs (1920) anhand bestimmter Eigenschaften als verschieden schwierig eingestuft, und zwar folgendermaßen:

1. Verwendung voller Seiten
2. Verwendung weniger diagonal geteilter Seiten
3. Verwendung aller diagonal geteilter Seiten
4. Drehen des Musters auf eine seiner Ecken
5. Weglassen der äußeren Rahmenlinie
6. Steigerung der verwendeten Würfel
7. Steigerung der Asymmetrie im Muster
8. Verwendung weniger verschiedener Farben im Muster

Kohs gab damit zugleich die Richtung für die vielen folgenden Untersuchungen vor. Beispielsweise bezog Thurstone (1944) das Verfahren in seine faktorenanalytische Studie mit ein und kam zu dem Schluss, dass der *Blockdesign*-Test auf dem von ihm definierten Faktor *Freiheit von Gestaltbindung* lädt, welcher in Kapitel 5.3 ausführlicher besprochen wird. Wie sich in den nächsten Abschnitten zeigen wird, lassen sich Kohs' Grundgedanken aber auch in jüngeren Studien, die dem zugrundeliegenden Konstrukt von Verfahren der Art *Blockdesign* vermehrt mit experimentellen Untersuchungsdesigns nachgehen, wiederfinden.

## 4.1 Lösungsprozesse und Mustereigenschaften

Royer und Weitzel (1977) gingen den von Kohs aufgestellten Schwierigkeitsfaktoren, anhand des *Block Design* im WAIS auf den Grund. Die Schritte 5 und 8 steigern laut Autoren die *perceptual cohesiveness* des Musters. Sie definierten diesen Stimulusfaktor durch die Anzahl an aneinander grenzenden gleichfarbigen Würfelkanten. Wird also eine Fläche als einheitlich wahrgenommen, obwohl sie sich aus mehreren Würfeln zusammensetzt, so muss an entsprechender Stelle ein mentaler Schnitt durch das Muster gezogen werden, was die Bearbeitungszeit erhöht. Wird das Muster in der Vorlage mit einem passenden Raster versehen, verkürzt sich die Bearbeitungszeit hingegen drastisch (Royer & Weitzel, 1997).

Den Lösungsvorgang beschreiben Royer und Weitzel (1977) anhand von vier mentalen Operationen:

- a) Verarbeitung des Muster
- b) Rekodierung des Musters in Einheiten, die mit den möglichen Alternativen der Würfelseiten übereinstimmen
- c) Größentransformation von der Vorlage auf das durch die Würfel nachgelegte Muster
- d) Vergleich des Ergebnisses mit der Vorlage

Dementsprechend gehen die Autoren davon aus, dass vor allem gewisse, mit der Wahrnehmungsorganisation verbundene, Eigenschaften des Musters, wie eben die *perceptual cohesiveness*, die Schwierigkeit der Aufgaben bestimmen. Die figurale Merkfähigkeit bezogen auf das gesamte Muster werde nur minimal gefordert, nämlich wenn der eigene Fortschritt mit der Vorlage verglichen wird. Royer (1977) bezieht neben der Ungewissheit durch die Fülle an möglichen Kombinationen zusätzlich die Redundanz im Muster als Parameter mit in seine Studie ein.

Schorr, Bower und Kiernan (1982) drehen die *perceptual cohesiveness* quasi um. Sie definieren die Anzahl an *edge cues*, also die Angrenzungen von roten mit weißen Würfelkanten, als Leichtigkeitsfaktor. Entsprechend ihrer Argumentation erleichtert dies das Projektieren eines Rasters bzw. die Ortung der Würfelkanten, was für die Identifikation der korrekten Würfelseite vonnöten ist und ohne zusätzliche Hinweise eine höhere mentale Anstrengung erfordert. Sie berücksichtigen dabei, anders als Royer und Weitzel (1977), alle vier Kanten eines Würfels gleichzeitig. Die Ergebnisse ihrer Studie zeigen, dass insbesondere die zweifarbiges Seiten umso schneller erkannt werden, je mehr *edge cues* für sie zur Verfügung stehen. Außerdem werden einfarbige Würfelseiten generell schneller als zweifarbig identifiziert. Sie begründen dies mit einem mentalen Rotationsaufwand, der für die zweifarbiges Würfelseiten erbracht werden muss. Schorr et al. (1982) konnten außerdem feststellen, dass *edge cues* für einfarbige Würfelseiten hauptsächlich zu Beginn des Nachlegens einen Vorteil bieten, nicht jedoch wenn das nachgelegte Muster bereits in fortgeschrittenem Maße vorliegt.

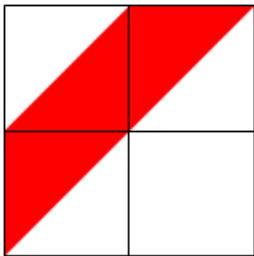
Rozencwajg und Corroyer (2001) gingen den Effekten struktureller Variablen der Muster nach, ebenfalls anhand der im WAIS zur Verfügung stehenden Würfelseiten. Allerdings bildeten in den Vorlagen dieser Studie lediglich die zweifarbiges Würfelseiten das eigentliche Muster. Die roten Würfelseiten wurden nicht mit in die Muster einbezogen. Die Autoren

gaben die speziell konstruierten Items 12- und 17-jährigen sowie erwachsenen Personen am Computer vor. Ziel war dabei die Erfassung von Effekten durch bestimmte Eigenschaften der abgebildeten Gestalten sowie der Bearbeitungsweise, welche durch die jeweiligen Mausklicks und Betrachtungszeiten erhoben wurde. Der Begriff *Gestalt* (vgl. Kapitel 5.1.3) bezieht sich hier auf die Wahrnehmung der kombinierten Musterelemente als kohärent und einheitlich. Rozencwajg und Corroyer (2001) konnten ihre Annahme bestätigen, dass Streifen<sup>1</sup> (Abbildung 2) deutlich schwieriger nachzulegen sind, als Rauten (Abbildung 3). Wenngleich mache Testpersonen keinerlei Probleme damit hatten, konnte der Effekt für alle getesteten Altersgruppen gefunden werden, wobei die 12-Jährigen die größten Schwierigkeiten damit zeigten. Sie legten die Streifen oft zu lang oder zu breit nach, sodass das Resultat nicht mehr die richtige Größe (3\*3 Würfelmatrix) hatte. Rauten hingegen scheinen den Testpersonen Informationen über die richtigen Proportionen des Musters zu liefern (Rozencwajg & Corroyer, 2001).

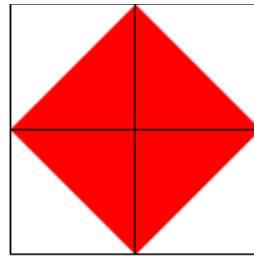
Die Autoren kommen zu dem Schluss, dass die Schwierigkeit des *Blockdesign* im Zergliedern des Musters in seine Elemente liegt und, dass die Testperson im Durchführen dieser kognitiven Operation durch Gestalten mit bestimmten Formen irregeleitet werden kann. Demnach ist das Nachlegen von Rauten deshalb einfacher als das Nachlegen von Streifen, weil die Achsen-Ausrichtung (horizontal-vertikal), mit der die Blöcke platziert werden müssen, den Achsen der Streifen zuwiderläuft, nicht jedoch denen der Diamantform. In beiden Fällen müssen die zweifarbigen Blöcke horizontal zusammengefügt werden. Die Streifen werden allerdings mit einer diagonalen Achse wahrgenommen. Rozencwajg und Corroyer verweisen hierzu auf eine mögliche Verbindung zur Feldunabhängigkeit (Das Konzept der Feldabhängigkeit/Feldunabhängigkeit wird in Kapitel 5.2 besprochen). Warum die Diamantform horizontal und nicht durch ihren Umriss auch diagonal, wie die Streifen, betrachtet wird, geht allerdings nicht eindeutig aus der Argumentation hervor. Die Tatsache, dass nicht alle Personen entsprechende Schwierigkeiten mit dem Nachlegen hatten, erklären die Autoren folgendermaßen: Entweder werden diese Personen nicht durch diesen Konflikt beeinträchtigt, wie es der Definition für Feldunabhängige entspricht, oder aber sie legen über das ganze Muster ein mentales Raster, das die Streifen zerteilt.

---

<sup>1</sup> Da die zweifarbigen Würfelseiten hier lediglich diagonal geteilt sind, und nicht wie im UT 10 auch andere Musterelemente zeigen können, lassen sich Streifen hier nur in diagonaler Form legen.



**Abbildung 2:** Beispiel eines möglichen Streifenmusters mit diagonal geteilten Würfelseiten



**Abbildung 3:** Beispiel eines möglichen Rautenmusters mit diagonal geteilten Würfelseiten

## 4.2 Bearbeitungsstrategien

Bei Untersuchungen mit dem *Blockdesign* und ähnlichen Verfahren, wurden oft zwei verschiedene Bearbeitungsweisen identifiziert. Die sich weitgehend deckenden Beobachtungen wurden meist mit den Begriffen *analytisch* und *synthetisch* (z.B. Schorr et al., 1982) aber auch *analytisch* und *global* (z.B. Beuscart-Zéphir & Beuscart, 1988) bezeichnet. Unter der analytischen Strategie wird gemeinhin ein Vorgehen verstanden, bei dem zunächst das vorliegende Muster mental in die jeweiligen Einheiten entsprechend der Würfelseiten segmentiert wird, und daraufhin die Würfel, einer nach dem anderen, an die entsprechende Stelle platziert werden. Bei der synthetischen bzw. globalen Herangehensweise hingegen wird das Muster nicht aufgegliedert, sondern als Ganzes wahrgenommen. Die Würfel werden an entsprechender Stelle so lange gedreht, bis sie in das Muster passen. Beim synthetischen Vorgehen liegt der Fokus auf der Gestalt des Vorlagemusters (Schorr et al., 1982). Zudem zeigte sich konsistent, dass erfolgreiche Testpersonen vorwiegend eine analytische Strategie verfolgen und damit vor allem schneller sind, als global vorgehende Personen (Beuscart-Zéphir & Beuscart, 1988; Rozencwajg & Corroyer, 2001; Schorr et al., 1982).

Schorr et al. (1982) bringen eine gewisse Rotationsleistung bei zweifarbigen Würfelseiten mit der analytischen Strategie in Zusammenhang. Einerseits muss die richtige Würfelseite im strukturierten Muster identifiziert werden, welche in vier verschiedenen Ausrichtungen vorliegen kann, und andererseits muss ein entsprechender Würfel ausgewählt und in die richtige Ausrichtung gebracht werden.

Wenngleich Schorr et al. (1982) von einer individuellen Tendenz zu einer der beiden Strategien ausgehen, so argumentieren sie dennoch, dass die Mustereigenschaften auffordernden Charakter für eine bestimmte Vorgehensweise haben können. Je stärker die Gestalt, also die Kohärenz, des abgebildeten Musters, desto eher wird eine synthetische Bearbeitung evoziert. Je weniger Informationen über die Kanten der jeweiligen Würfelseiten im Muster enthalten sind (*edge cues*), desto eher ist eine analytische Segmentierung des Musters vonnöten. Die Untersuchung wurde mit Studenten durchgeführt, denen beinahe allen eine vorwiegend analytische Vorgehensweise attestiert wurde. Daher kann die Aussagekraft dieser Schlussfolgerungen in Frage gestellt werden. Schorr et al. argumentieren allerdings weiter, dass besonders zu Beginn des Nachlegens eher analytisch vorgegangen wird, wohingegen später das bereits teilweise zusammengesetzte Muster stärkere Gestalt aufweist und daher – auch bei vorwiegend analytisch vorgehenden Personen – synthetisches Vorgehen begünstigt.

Rozencwajg und Corroyer (2001) fanden ebenfalls einen Zusammenhang zwischen Mustereigenschaften und Bearbeitungsweise. Besonders die schwierigeren, gestreiften Muster ihrer Studie evozierten ein globales Vorgehen.

Beuscart-Zéphir und Beuscart (1988) sprechen von drei individuellen Profilen: analytisch vorgehenden Personen mit hoher Erfolgsrate, global vorgehenden Personen und Personen, welche quasi dazwischen liegen. Letztere versuchen zunächst eine analytische Strategie, wobei sie allerdings häufig versagen und folglich auf ein globales Vorgehen zurückfallen.

Die Schlussfolgerungen von Schorr et al. (1982) sowie Beuscart-Zéphir und Beuscart (1988) stehen im Einklang mit Rozencwajg und Corroyer (2001). Die Autoren differenzieren drei Strategien. Zusätzlich zu einer globalen und einer analytischen Vorgehensweise definieren sie eine weitere mit der Bezeichnung „synthetisch“. Demnach gehen manche Personen kombiniert analytisch und global vor, was sich in dieser Studie als die erfolgreichste Strategie herausstellte. Für das analytische Vorgehen sind laut den Autoren, ähnlich Schorr et al. (1982), zwei kognitive Operationen charakteristisch, nämlich die Segmentierung des Musters sowie die mentale Rotation der zweifärbigen Würfelseiten. Die Würfel werden Reihe für Reihe, unabhängig vom Muster, platziert. Personen, die global vorgehen, führen diese Operationen nicht durch, sondern passen die Würfel nach und nach durch Versuch und Irrtum dem Muster, entsprechend der abgebildeten Gestalten, an. Die Ausrichtung der zweifärbigen Würfelseiten geschieht durch das drehen in der Hand und versuchsweise Platzierung. Die

Autoren zeigen auf, dass im Extremfall eine Person durch die globale Vorgehensweise das Muster nicht fertig nachlegen kann. Beim „synthetischen“ Vorgehen sind wiederum Segmentierung und mentale Rotation beteiligt, jedoch werden zunächst die im Muster enthaltenen Gestalten nachgelegt.

In ihrer Vorgängerstudie konnte Rozencwajg (1991) zudem zeigen, dass ein Zusammenhang zwischen der Vorgehensweise und der Ausprägung in Feldabhängigkeit/Feldunabhängigkeit (siehe Kapitel 5.2) besteht. Feldabhängige gehen laut Rozencwajg global vor, Feldunabhängige kombiniert global und analytisch und die dazwischen liegenden Personen bearbeiten die Aufgaben analytisch. Auch Kurian, Sharma und Santhakumari (1987) kamen zu ähnlichen Ergebnissen, wonach feldabhängige Personen bei *Blockdesign*-Aufgaben eher global vorgehen, indem sie die Würfel an fixer Position liegend öfters drehen, währendfeldunabhängige die einzelnen Teile des vorher segmentierten Musters nachlegen, anstatt das Muster als Ganzes anzugehen. Rozencwajg (1991) argumentiert, dass Personen im mittleren Feldabhängigkeit/Feldunabhängigkeits-Bereich die Muster Würfel für Würfel analysieren während Feldunabhängige sich stärker vom Musterdesign leiten lassen können. Rozencwajg und Corroyer (2001) konnten zudem zeigen, dass 17-jährige und erwachsene Testpersonen *Blockdesign*-Aufgaben zumeist auf diese Weise lösen, während 12-Jährige meist global vorgehen. Dies deutet, zusammen mit den Ergebnissen von Schorr et al. (1982) auf einen Entwicklungsaspekt der im *Blockdesign* erfassten Fähigkeit hin.

## **5. Zur kognitiven Organisation visueller Wahrnehmungsinformationen**

Um die Lösungsprozesse nachzuvollziehen, mit denen die beschriebenen Vorlagen segmentiert und wieder zusammengesetzt werden, ist unter anderem das Verständnis der zugrundeliegenden Phänomene der visuellen Wahrnehmungsorganisation vonnöten. Hiermit sind nicht die biophysiologischen Voraussetzungen und Prozesse der visuellen Reizverarbeitung gemeint, auch nicht das (Wieder)Erkennen und Identifizieren bestimmter Objekte, dies greift zu weit. Vielmehr ist die Organisation der Fülle, der durch das Auge erfassten Informationen, zu einzelnen Einheiten von Interesse. Maderthaner (2008) beschreibt den kognitiven Ablauf bei der Verarbeitung visueller Information als Strukturierung der

visuellen Information, sodass Elemente bzw. Objekteinheiten abgegrenzt und identifiziert werden können.

Die Frage, wie diese Strukturierung von statthen geht, ist von entsprechender Bedeutsamkeit, was in vollem Umfang erstmals von Wertheimer (1923) erfasst und diskutiert wurde. Er gilt als einer der Gründer der *Gestaltpsychologie*, eine von vielen Wissenschaftsrichtungen, die sich mit der Wahrnehmung von Objekten beschäftigt. Die von der Gestaltpsychologie aufgestellten Wahrnehmungsprinzipien können für die vorliegende Arbeit Hinweise auf relevante Bildinformationen geben, da sie im Besonderen auf die Beziehung von Reizkombinationen und Wahrnehmung eingehen. Daher sollen diese Prinzipien in Bezug auf deren Bedeutung für die Gliederung mehr oder weniger einfacher Muster einer genaueren Betrachtung unterzogen werden.

Ein bereits konkreterer Zusammenhang mit den für den UT 10 relevanten kognitiven Prozessen kann – wie bereits in den vorangegangenen Abschnitten angeführt – für das Konzept der Feldabhängigkeit/Feldunabhängigkeit angenommen werden. Dies beschreibt, grob zusammengefasst, eine gewisse Flexibilität im Umgang mit Bildinformationen. Die in diesem Bereich entwickelten Definitionen und die damit zusammenhängenden Untersuchungen zeigen große Parallelen zum generellen Verständnis der dem UT 10 zugrunde liegenden Fähigkeit auf. Entsprechend Hergovich (1992) muss sogar davon ausgegangen werden, dass Feldabhängigkeit/Feldunabhängigkeit und die im UT 10 gemessene Fähigkeit ein und dasselbe sind. Auch zuvor Witkin (1975, zitiert nach Sternberg & Grigorenko, 1997) behauptete, dass Feldunabhängigkeit mit der geforderten Fähigkeit im *Block Design* ident ist. Goodenough und Karp (1961) fanden ebenfalls einen Zusammenhang mit dem *Block Design*.

Die Auffassung von Feldabhängigkeit/Feldunabhängigkeit in der Psychologie wird jedoch sehr kontrovers diskutiert, weshalb die konkrete Bestimmung kognitiver Operationen durch diesen sehr breiten Begriff nicht eindeutig, vor allem aber nicht direkt, gelingt. Das Verständnis dieser Begrifflichkeit soll daher unter Berücksichtigung dessen zeitlichen Wandels etwas genauer betrachtet werden. Die operanten Definitionen für die damit verbundenen psychologischen Tests sollen eine direktere Verbindung zum UT 10 erlauben.

Mit dem Konzept der Feldabhängigkeit/Feldunabhängigkeit verwandt oder gar ident ist das Konstrukt der *Freiheit von Gestaltbindung*. Es integriert allerdings die Bedeutung von

Gestaltgesetzen stärker in die Bestimmung der mit dem *Blockdesign* verbundenen kognitiven Prozesse. Basis dieses Konzepts stellen diverse faktorenanalytische Studien dar, die einer differenzierteren Betrachtung von figuralem Material mit Gestalt-Charakter dienlich sind und zu einem besseren Verständnis der Gemeinsamkeiten und Unterschiede solcher Verfahren bezüglich der erfassten Konstrukte beitragen können.

## 5.1 Die Gestaltpsychologie und ihre Gesetze

Die Anfänge der Gestaltpsychologie liegen bereits um die hundert Jahre zurück. Aufbauend auf den Überlegungen von Max Wertheimer wurden von ihm und Kollegen, wie beispielsweise Wolfgang Köhler und Kurt Koffka, sogenannte *Gestaltfaktoren* formuliert (King & Wertheimer, 2005). Sie beschreiben Eigenschaften von Reizen, welche zur Zusammenfassung einzelner Wahrnehmungselemente zu Einheiten führen (Maderthaner, 2008), bzw. etwas genereller formuliert, welche Wahrnehmungen sich ergeben, wenn gewisse Reizbedingungen angetroffen werden (Goldstein, 2002). Sowohl die Bezeichnung *Gestaltfaktoren* als auch die sehr gängige Formulierung *Gestaltgesetze* greifen jedoch etwas zu weit, da ihre Gültigkeit nach heutigem Kenntnisstand eher der von Faustregeln oder Heuristiken entspricht (z.B. Goldstein, 2010). Zudem handelt es sich um *ceteris paribus* Regeln. Ihre Gültigkeit kann nur dann mit einiger Sicherheit angenommen werden, wenn alle anderen Bedingungen gleich sind – also kein konkurrierendes Prinzip gleichzeitig im Wahrnehmungsfeld aufkommt (Wertheimer, 1923).

Heute gilt die Schule der Gestaltpsychologie als ganzheitliche Wahrnehmungstheorie als gescheitert, doch wird sie als sehr erfolgreicher Ansatz anerkannt und übt nach wie vor einen großen Einfluss auf die Wahrnehmungspsychologie aus (Goldstein, 2010). Zahlreiche Beispiele dafür, dass im Gebiet der Gestaltpsychologie auch aktuell noch Studien durchgeführt werden, die einerseits zum Verständnis der bereits formulierten Prinzipien beitragen, oder diese ergänzen, und andererseits neue Prinzipien hervorbringen, finden sich bei Wagemans et al. (2012). „Gestalt ideas are making a comeback in the guise of dynamic connectionist theories“ (Palmer, 1999, S. 53). So wurden beispielsweise in jüngeren Untersuchungen vermehrt komplexere Reize vorgegeben, die der natürlichen Umwelt eher entsprechen (Wagemans et al., 2012).

Dass die Treffsicherheit der mittels Gestaltgesetzen erstellten Vorhersagen in den meisten Untersuchungen sich als sehr stark an das speziell dafür entwickelte Anschauungsmaterial gebunden erwies, ist eine der am häufigsten angeführten Kritiken an der Gestalttheorie (z.B. Goldstein, 2002, 2010; Gordon, 2004). Für die vorliegende Arbeit wird dieser gerechtfertigten Beanstandung weniger Gewicht beigegeben, weil das Material des Untertests *Analysieren und Synthetisieren* in Qualität und Vorgabe eher diesen simpel strukturierten, zweidimensionalen und meist zweifarbigen Vorgaben in den von den Gestaltpsychologen durchgeführten Untersuchungen entspricht, als alltäglichen Umweltreizen.

Aus diesem Grund sollen einige der Prinzipien hier erwähnt werden. Sie werden in der Regel anhand ihrer Funktionen in zwei Gruppen gegliedert:

- Prinzipien zur Gruppierung einzelner Wahrnehmungselemente und
- Prinzipien zur Differenzierung zwischen Figur und Hintergrund.

Die Gruppierungsprinzipien bestimmen dabei die Elemente der Wahrnehmung und die Figur-Grund-Relationen bestimmen deren Interpretation bezüglich Form und Lage (Wagemans et al., 2012). In Betracht der Materialeigenschaften des UT 10 sind allerdings nur die Figur-Grund-Relationen von Bedeutung, und zwar jene, die sich auf zweidimensionale und statische Reize anwenden lassen. Aufbauend auf die von Wertheimer 1923 erstmals aufgestellten Faktoren lässt sich, nach heutigem Stand der Forschung und weitgehender Lehrmeinung, folgender Überblick über die für diese Arbeit relevanten Prinzipien der Figur-Grund-Trennung geben (vgl. Bruce, Green & Georgeson, 2004; Goldstein, 2002, 2010; Gordon, 2004; Palmer, 1999; Wagemans et al., 2012).

### 5.1.1 Prinzipien der Figur-Grund-Trennung

In der Figur-Grund-Trennung liegt die Fähigkeit, Gegenstände voneinander getrennt wahrzunehmen (Goldstein, 2002). Die Frage, wie diese Trennung zustande kommt, wird im Rahmen der Figur-Grund-Relationen diskutiert, „worunter die Charakteristika von Linien, Konturen und Flächen verstanden werden, die eine räumliche Schichtung nach Vordergrund und Hintergrund erlauben“ (Maderthaner, 2008, S. 146). Zwei aneinander liegende Regionen im Sichtfeld, die sich eine Grenze teilen, führen üblicher Weise zur Segmentierung in Figur und Hintergrund. Dabei wird die Grenzlinie als der Figur zugehörig empfunden (*border ownership*) und der Hintergrund als, ohne Grenze und Form, sich hinter der Figur

fortsetzend.<sup>[S1]</sup> Die Figur wird auch eher als gestalthaft empfunden und ist eindrücklicher als der Hintergrund (Wertheimer, 1923).

### Konvexität versus Konkavität

*Konvexe Formen werden eher als Figur wahrgenommen. Konkave Formen werden eher als Hintergrund wahrgenommen.*

Dies trifft in dieser allgemeinen Formulierung vorrangig auf zweifarbige Darstellungen mit mehr als einer Ein- bzw. Ausbuchtung zu. Bei mehreren verschiedenfarbigen Flächen spielt die Färbigkeit eine stärkere Rolle, so werden konkave Formen insbesondere dann eher als Hintergrund wahrgenommen, wenn diese dieselbe Farbe haben. Die Färbigkeit der konvexen Formen ist hingegen irrelevant (Peterson & Salvagio, 2008).

### Symmetrie

*Symmetrische Elemente werden eher als Figur wahrgenommen als nicht symmetrische* (Peterson & Gibson, 1994).

Insbesondere konvexe, symmetrische Formen haben also starken Figurcharakter.

### Größe

*Kleinere Bereiche werden eher als Figur wahrgenommen.*

Ist ein Bild in zwei gleichartige aber unterschiedlich große Bereiche geteilt, so wird in der Regel der kleinere Bereich als Figur vor einem größeren Hintergrund interpretiert (Goldstein, 2010).

### lower region

*Der untere Teil einer in zwei Bereiche getrennten Reizanordnung wird eher als Figur wahrgenommen als der obere Teil* (Vecera, Vogel & Woodman, 2002).

Vecera et al. (2002) führen als ökologische Erklärung für dieses Prinzip an, dass in der natürlichen Umwelt der Bereich unterhalb der Horizontlinie bzw. Objekte innerhalb dieses Bereichs in der Regel näher beim Betrachter liegen als jene oberhalb der Horizontlinie. Die Autoren schließen darauf, dass die Position im Bezug zur Horizontlinie auch einen Tiefenhinweis darstellen kann (Vecera et al., 2002; Vecera & Palmer, 2006).

### **top-bottom polarity**

*Formen, die unten breiter sind als oben, werden eher als Figur wahrgenommen, als Formen, die oben breiter sind als unten* (Hulleman & Humphreys, 2004).

Wagemans et al. (2012) argumentieren, dass hier wie bei *lower region* die Schwerkraft als ökologische Erklärung dient.

Was den bisherigen Ausführungen zugrunde liegt, ist das Selbstverständnis, mit dem von den zu gruppierenden und zu trennenden Elementen ausgegangen wird. Wie die Wahrnehmung von diesen Elementen, wie Punkte, Striche und Flächen, überhaupt erst zustande kommt, erklärt Palmer & Rock (1994) mit dem Prinzip der *uniform connectedness*, welches hier der Vollständigkeit halber erwähnt wird, aber nicht genauer besprochen werden soll.

### **5.1.2 Wahrnehmungsorganisation im Kindesalter**

In mehreren Untersuchungen zeigte sich, dass die Wahrnehmung Neugeborener nicht von Anfang an allen statischen Gestaltgesetzen folgt (z.B. Hofsten & Spelke, 1985; Spelke, 1990; Spelke, Breinlinger, Jacobson & Phillips, 1993; Spelke, Hofsten & Kestenbaum, 1989). Laut der genannten Studien tragen beispielsweise Ähnlichkeit und Kontinuität bei Neugeborenen unter einem halben Jahr nicht dazu bei, nebeneinanderliegende Objekte zu trennen. Vielmehr werden diese und andere Prinzipien erst im Laufe des ersten Lebensjahres zunehmend angewandt. Im Allgemeinen kann aus diesen und anderen Studien (z.B. Needham & Baillargeon, 1997; Quinn & Bhatt, 2001) abgeleitet werden, dass die Entwicklung der Wahrnehmungsorganisation anhand von Gestaltprinzipien ungefähr ab dem 8. Lebensmonat abgeschlossen ist.

### **5.1.3 Die gute Gestalt**

Das Prinzip der Prägnanz wurde von Wertheimer (1923) als ein allen Gestaltgesetzen zugrunde liegendes Metaprinzip formuliert. Es besagt, dass die wahrgenommenen Objekte stets die einfachste unter den gegebenen Bedingungen mögliche Form annehmen. Eine in diesem Sinne *gute Gestalt* ist möglichst symmetrisch sowie vollständig und steht in der Regel prägnant vor dem Hintergrund heraus. Eine gute Gestalt ist kohärent und hat den Charakter einer Einheit. Köhler (1933) zufolge bedeutet das Wort „Gestalt“ in der Gestalttheorie “ein ausgesondertes Ganzes” (S. 122). Dies macht deutlich, dass eine gute Gestalt mit durch die Geltung von Figur-Grund-Prinzipien bestimmt wird.

Köhler (1933) argumentiert, dass für eine anschauliche Form jene Reizkonstellationen verantwortlich sind, die für den Formcharakter und die Aussonderung dieses Ganzen maßgebend sind. „Eine Formbeschaffenheit ändert sich hingegen nicht wesentlich, wenn man Helligkeit, Farbe, Grösse und Ort innerhalb weiter Grenzen nach bestimmten Gesetzen variiert“ (S. 135). Dies betont einerseits die Rolle der Beziehungen zwischen den einzelnen figurbildenden Informationen und zeigt andererseits, dass sich durch die Anwesenheit vieler bzw. anderer Reize die Relevanz der einzelnen Prinzipien verschiebt. Die Wahrnehmung von Objekten ist also viel komplexer, als die Aufzählung einzelner Prinzipien vermuten lässt. Es handelt sich dabei um einen dynamischen Prozess (King & Wertheimer, 2005).

Dieser Prozess kann bei gewissen kognitiven Leistungen eine Rolle spielen. Clement und Varnadoe (1967) konnten in ihrer Untersuchung zur wahrnehmungsbezogenen Diskriminationsfähigkeit feststellen, dass zwei jeweils gute Gestalten deutlich besser voneinander unterschieden werden können als zwei „schlechte“ Figuren. Garner und Sutliff (1974) wiederholten diese Studie mit übereinstimmendem Resultat. Dabei wurde auch die jeweilige Reaktionszeit miterfasst. Aus den Ergebnissen kann geschlossen werden, dass gute Muster schneller wahrgenommen bzw. enkodiert werden, als schlechte. Dies konnte auch durch Bell und Handel (1976) bestätigt werden. Gute Gestalten werden außerdem besser erinnert (Checkosky & Whitlock, 1973). Attnave (1955) ging dem Einfluss auf die Gedächtnisleistung für Punkt-Muster in vordefinierten Matrizen nach. Es zeigte sich, dass bei konstant gehaltener Größe der Matrix, für symmetrische Muster eine bessere Reproduktionsleistung festgestellt werden kann. Allerdings gilt: je größer das Muster, desto schlechter die Merkleistung. Kleine asymmetrische Muster werden besser gemerkt als sehr große, auch wenn diese mehrfach symmetrisch sind.

Was eine gute Gestalt ist, scheint intuitiv und relativ einleuchtend. Die Güte einer Gestalt allgemein zu beurteilen und in einen vergleichbaren Wert zu fassen stellt hingegen eine größere Herausforderung dar. Ein aus der Gestaltpsychologie hervorgehender Ansatz zur Messung der Prägnanz einer abgebildeten Figur stellt die *Structural Information Theory* (SIT) dar.

### 5.1.4 Operationalisierung von guter Gestalt mittels SIT

Die *Structural Information Theory (SIT)* liefert eine Methode zur Beschreibung von Formen, mit der es auch möglich ist, die Qualität einer Form im Sinne einer guten Gestalt bzw. Prägnanz zu erfassen und zu vergleichen. Dabei werden die gesamten Bildinformationen zu einem repräsentativen Wert zusammengefasst. Die Methode wurde von Leeuwenberg (1969, 1971) ins Leben gerufen und von ihm zusammen mit seinen Kollegen (z.B. Buffart & Leeuwenberg, 1983; Buffart, Leeuwenberg & Restle, 1981; Restle, 1979) weiter entwickelt. Die *Structural Information Theory* gilt für die menschliche Interpretation von physikalischen Objekten oder Ereignissen, soll jedoch nicht eine globale Theorie über den Wahrnehmungsprozess darstellen (Buffart & Leeuwenberg, 1983). Sie beschreibt lediglich die Beziehung zwischen Reizbeschaffenheit und dessen möglichen Formrepräsentationen, nicht jedoch *wie* es vom Input zum Output kommt.

Idealer Weise werden zunächst alle laut Bildinformation möglichen Formen (was meist unmöglich ist und sich auf die naheliegendsten beschränkt) abgeleitet und daraufhin folgendermaßen beschrieben und verglichen:

Die Beschreibung einer Form geschieht anhand von sogenannten *codes*, welche ausreichen um die Form zu reproduzieren. Sie beschreiben unter anderem gerade Linien und verschiedene Winkel. Zuerst wird von der Form ein *primitive code* abgeleitet, indem der Verlauf der Form durch eine Aneinanderreihung von Symbolen für Linien, Winkel und Ähnliches. beschrieben wird. Dieser Code wird dann anhand bestimmter Regeln und Funktionen so weit wie möglich zu *reduced codes* vereinfacht. Die Verbindung zur Prägnanz schafft der sogenannte *information load* (I). Der *information load* ergibt sich aus der Anzahl an Parametern und erfasst damit die Komplexität der generierten Codes. Er soll dazu dienen, den besten, also den einfachsten, Code zu identifizieren (Restle, 1979). Dieser *minimum code* beschreibt die Form, die bei gegebener Bildinformation laut *SIT* von Probanden am häufigsten wahrgenommen wird. Beispielsweise konnten Buffart et al. (1981) zeigen, dass anhand des *information loads* die Wahrnehmungsorganisation in mehrdeutigen Situationen erklärt werden kann.

Palmer (1999) fasst die Stärken und Schwächen der *Structural Information Theory* zusammen. Zu den Stärken zählen demnach die Möglichkeit, Symmetrie und Regelmäßigkeit zu erfassen, sowie die plausible Erklärung vieler Wahrnehmungssphänomene. Allerdings ist

die *SIT* nur für idealisierte Umrisszeichnungen geeignet. Eine Anwendung auf schattierte oder grau skalierte Bilder ohne klare Linien ist nicht möglich. Zudem gibt es keine expliziten Codes für andere Verläufe als die aus simplen Linien oder Kurven bestehenden. Zum Beispiel werden weder Konkavität noch Konvexität berücksichtigt. Vor allem ist die Anwendung auf zweidimensionale Darstellungen zwar vergleichsweise simpel, doch bleibt bei der Umlegung auf dreidimensionale Darstellungen offen, wie hier von einer 2-D Darstellung auf die 3-D Projektion geschlossen wird. Die *SIT* erhebt zwar nicht den Anspruch, als Prozessmodell der visuellen Wahrnehmung zu dienen, hier bleibt jedoch eine große Lücke offen. Anwendungsnähere Kritik betrifft vor allem die hohe Sensitivität gegenüber geringen Abweichungen und die unrealistische Forderung, *alle möglichen* Codes zu erstellen und zu vergleichen.

Im Anschluss an diese kritische Betrachtung gilt es nochmals zu erwähnen, dass die *SIT* sich betreffend zweidimensionaler Abbildungen simpler Formumrisse als eine gute Operationalisierungsmöglichkeit gezeigt hat, um Formen zu beschreiben und deren Komplexität zu erfassen.

## 5.2 Das Konzept Feldabhängigkeit/Feldunabhängigkeit

Feldabhängigkeit und Feldunabhängigkeit sind in der Psychologie heute weithin bekannte Begriffe. Deren Auffassung ist aber nicht immer einheitlich. Das Verständnis von Feldabhängigkeit bzw. Feldunabhängigkeit hat sich seit der Entstehung des Konzepts mehrere Male geändert und führte in den vergangenen Jahrzehnten zu kontroversen Diskussionen und regem Forschungsaufwand (einen kurzen Überblick bieten z.B. Evans, Richardson & Waring, 2013).

Sowohl die Entstehung als auch die Veränderung des Konzepts war stets eng an die Entwicklung von bestimmten Verfahren und die dadurch gewonnenen Erkenntnisse gebunden. Die Gemeinsamkeiten dieser Tests mit dem UT 10, sowohl bezüglich des Materials als auch der Definition des zu erfassenden Konstrukts, liefern für diese Arbeit nützliche Hinweise für das Verständnis der kognitiven Prozesse bei ähnlichen wahrnehmungsbezogenen Aufgaben. Insbesondere die Ergebnisse von Hergovich (1992; Kapitel 3.3) machen die Relevanz dieses Konzepts für diese Arbeit deutlich. Deshalb soll für Feldabhängigkeit/Feldunabhängigkeit (FDI) als nächstes ein kurzer historischer Überblick über die Definition und Operationalisierung sowie mögliche Kritikpunkte gegeben werden.

### **5.2.1 Historischer Überblick von Feldabhängigkeit/ Feldunabhängigkeit**

Seinen Ursprung hat das Konzept der Feldabhängigkeit /Feldunabhängigkeit in Herman A. Witkins Untersuchungen zum Zusammenwirken bzw. zum Konkurrieren visueller und vestibulärer Informationen für die Orientierung im Raum (Ash & Witkin, 1948a, 1948b; Witkin & Ash, 1948a, 1948b). Bei den Experimenten mit schwenkbarem Rahmen (*rod-and-frame test*, RFT) oder schwenkbarem Raum und kippbarem Stuhl (auf dem die Probanden saßen; *tilting-room—tilting-chair test*, TRTC) wurde die Beobachtung gemacht, dass sich Personen stark darin unterscheiden, inwieweit sie sich eher an den visuellen oder den vestibulären Informationen orientieren.

Es wurden erstmals die Begriffe *field dependence* (FD) und *field independence* (FI) eingeführt. Als feldabhängig galt eine Person dann, wenn sie sich bei der senkrechten Ausrichtung eines Objektes oder ihres eigenen Körpers vorrangig an am visuellen Umfeld orientierte und eine davon unabhängige Wahrnehmung nicht gelang. Als feldunabhängig bezeichnete Personen hingegen orientierten sich eher an vestibulären Empfindungen und erzielten bessere Ergebnisse. In seinen Anfängen wurde Feldabhängigkeit/Feldunabhängigkeit somit als relative Dominanz visueller oder vestibulärer Sinneseindrücke verstanden (Witkin et al., 1962).

Während die Rolle vestibulärer Sinneseindrücke mit der Zeit für das Konzept stark an Bedeutung verlor, wurde die Hypothese aufgestellt, dass Unterschiede in der „ability to overcome an embedding context“ (Witkin, Dyk, Faterson, Goodenough & Karp, 1962, S. 43) für die konsistenten Ergebnisse verantwortlich sind.

Bei den sogenannten *Gottschaldt-Figuren* (Gottschaldt, 1929) wird eine einfache, abstrakte, aus geometrisch angeordneten Linien bestehende Figur vorgegeben, die in einer komplexeren, abstrakten und ebenfalls aus solchen Linien bestehenden Figur eingebettet entdeckt werden soll. Die versteckte Figur hat dabei immer dieselbe Größe und Stellung. Es geht also um das Herauslösen eines Teils aus einem Ganzen. Dieterich (1975) meint dazu: „Die versteckte Figur [...] kann aber nur gefunden werden, wenn die spontane nach Gestaltgesetzen (z.B. Geschlossenheit, Nähe, durchgehende Linie) entstehende Organisation der komplexen Figur aufgelöst wird. Oder anders formuliert: das vorherrschende Feld muß umstrukturiert werden“ (S.86).

Dasselbe Prinzip wendet Witkin in seinem *Embedded Figures Test* (EFT; Witkin, Oltman, Raskin & Karp, 1971) an, wobei in die komplexen Figuren auch Farbelemente mit einbezogen werden. Witkin et al. (1971) bezeichnen den Lösungsprozess als *disembedding* und argumentieren, dass relativ feldabhängige Personen der Organisation des dargebotenen Feldes folgen, während die feldunabhängigeren Personen die Organisation des Feldes überwinden und aufbrechen können, um die gesuchte Komponente zu finden.

Durch weitere Beobachtungen bei und von wahrnehmungsbezogenen Untersuchungen kamen Witkin et al. (1962) zu dem Schluss, dass es sich bei Feldabhängigkeit/Feldunabhängigkeit um eine zugrundeliegende Disposition einer Person mit verschiedenen Ausprägungsgraden handelt, welche in Zusammenhang mit einer Vielzahl an Persönlichkeitscharakteristika und psychologischen Funktionen steht. Sie sprechen von *psychological differentiation* als entwicklungsbedingte Fähigkeit, die Umwelt zu Analysieren und Erfahrungen zu strukturieren.

The person who experiences in articulated fashion has the ability to perceive items as discrete from their backgrounds, or to reorganize a field, when the field is organized; and to impose structure on a field, and so perceive it as organized, when the field has relatively little inherent structure. In this view the ability to analyze experience and the ability to structure experience are both aspects of increasing articulation. (S. 14)

Auch in anderen Wahrnehmungsbereichen scheint Feldabhängigkeit/Feldunabhängigkeit eine Rolle zu spielen. So beziehen sich Witkin et al. (1962) auf mehrere Untersuchungen zur *Wahrnehmungskonstanz* [s2] bei denen sowohl eine *globale* als auch eine *analytische* Betrachtungsweise gefordert wurde. Die Autoren schließen aus den Ergebnissen, dass Feldunabhängige sowohl eine analytische als auch eine globale Herangehensweise annehmen können, während Feldabhängige auf die globale Bearbeitungsweise beschränkt sind. Sie generalisieren diesen Gedanken auf die *psychological differentiation*, wonach eine hohe Ausprägung hier bedeutet, dass die Person auf einem hohen Level zu dieser Art von Differenzierung fähig ist, diese jedoch nicht ständig anwendet und auch auf einem niedrigeren Level ausübt. Ob die höchst mögliche Differenzierung genutzt wird oder nicht, hängt von der Motivation und / oder den Anforderungen der jeweiligen Situation ab.

Es wurde des Weiteren in Betracht gezogen, dass die Fähigkeit, Elemente aus dem Kontext zu nehmen und in neue Zusammenhänge zu bringen, nicht nur in wahrnehmungsbezogenen

Situationen zum Tragen kommt, sondern auch beim Problemlösen (Witkin et al., 1962). Witkin et al. (1962) berichten von einer Reihe von Untersuchungen in den 50er-Jahren, die sich mit dem Zusammenhang zwischen wahrnehmungsbezogener Feldunabhängigkeit und der Leistung in standardmäßigen Intelligenz- und Leistungstests auseinandersetzen. Zusammen mit den Ergebnissen der berechneten Faktorenanalyse, die hohe Zusammenhänge mit bestimmten, aber nicht allen, intelligenzbezogenen Leistungstests zeigte, sehen die Autoren den Beweis dafür, dass es sich bei Feldunabhängigkeit um einen kognitiven Stil und nicht um einen grundlegenden Intelligenzfaktor handelt. Sie orten den gemeinsamen Nenner in der Fähigkeit, Kontextgebundenheit zu überwinden, und sprechen von *analytical functioning*.

Feldabhängigkeit und Feldunabhängigkeit werden als zwei Pole einer Dimension, des kognitiven Stils, gesehen. Demnach bergen beide Ausprägungsrichtungen ihre Vor- und Nachteile und sind voneinander nicht als besser oder schlechter unterscheidbar. Feldunabhängige zeigen Stärken zum Beispiel im *restructuring*, Feldabhängige hingegen in Bereichen der sozialen Kompetenz (Witkin et al., 1962).

Das komplexe Konstruktgebilde rund um die *psychological differentiation* und Feldabhängigkeit/Feldunabhängigkeit fand in den vergangenen Jahrzehnten nicht nur Anhänger sondern auch Kritiker (z.B. Kogan, 1980, zitiert nach Hergovich, 1992; Kogan & Block, 1991; McKenna, 1984). Einer der bedeutendsten Streitpunkte betrifft dabei die Dimensionalität von Feldabhängigkeit/Feldunabhängigkeit. Bevor hierauf genauer eingegangen wird, soll sich dem Entwicklungsaspekt gewidmet werden.

### 5.2.2 Entwicklungsaspekte der Feldabhängigkeit/Feldunabhängigkeit

Schon 1954 konnten Witkin und seine Mitarbeiter feststellen, dass besonders junge Kinder zu einer stärker feldabhängigen Wahrnehmung tendieren und mit zunehmendem Alter eine mehr feldunabhängige Wahrnehmung entwickeln (Witkin, Lewis, Hertzman, Machover, Meissner & Wapner, 1972). Auch Witkin, Goodenough und Karp (1967) entdeckten in ihrer Langzeitstudie einerseits einen bedeutenden Zusammenhang zwischen der Leistung in Feldunabhängigkeitstests und Alter sowie andererseits eine hohe relative Stabilität der individuellen Ausprägung. Bei Untersuchungen mit dem EFT, dem RFT und dem TRTC bei 8- bis 24-Jährigen zeigte sich ein durchgehender Anstieg an Feldunabhängigkeit zwischen 8 und 15 Jahren, wobei die Anstiegsrate gegen Ende des Zeitraums hin zunehmend schwächer

wurde. Kinder als Individuen behielten ihre Position bezüglich Feldunabhängigkeit im Vergleich zu Gleichaltrigen recht stabil bei, während alle Kinder gesamt sich kontinuierlich mehr in Richtung Feldunabhängigkeit entwickelten. Auch nachfolgende Untersuchungen zeigten entsprechende Ergebnisse. Witkins (1976) Arbeit ergab, dass insbesondere die Entwicklung zwischen 5 und 8 Jahren schneller verläuft als die nachfolgende. Köstlin-Gloge (1978) setzt das Alter, ab dem die Entwicklung der Feldunabhängigkeit weit genug fortgeschritten ist, um von einer Wahlfreiheit zwischen feldabhängigem und feldunabhängigem Denken sprechen zu können, mit 15 Jahren an.

### **5.2.3 Zur Dimensionalität von Feldabhängigkeit/Feldunabhängigkeit**

Im Widerspruch zur behaupteten Bipolarität von Feldabhängigkeit/Feldunabhängigkeit stehen zunächst einmal die voneinander verschiedenen Eigenschaften, die Feldabhängigkeit und Feldunabhängigkeit zugesprochen werden. Lägen beide auf einer Dimension, so dürfte es sich nur um stärkere oder schwächere Ausprägungen derselben Eigenschaften handeln. Des Weiteren sind die Tests für Feldunabhängigkeit und die für Feldabhängigkeit voneinander verschieden. Lägen beide auf einer Dimension, so müsste man die Tests im Sinne der modernen Testtheorie (siehe Kapitel 7) zu einem Test zusammenfassen können. Genau genommen müsste es sogar ausreichen, nur eine Art von Aufgaben vorzugeben, um beide Pole zu erfassen (Hergovich, 1992). Jeder Person müsste ein gewisser Grad an Feldabhängigkeit/Feldunabhängigkeit zugeordnet werden können, der relative stabil ist. Man hätte also nicht die Möglichkeit, zwischen den beiden Stilen zu wechseln. Genau das können aber zumindest Feldunabhängige anscheinend machen (Kogan, 1980, zitiert nach Hergovich, 1992).

Wenn man den Entwicklungsaspekt berücksichtigt wonach Kinder feldabhängiger sind, so müssten diese entsprechend Witkin über eine höhere soziale Intelligenz als Erwachsene bzw. feldunabhängige Personen verfügen. Dies scheint genauso wenig plausibel, wie dass höhere Empathie erst durch eine mangelnde Differenzierungsfähigkeit zustande kommt (Kogan & Block, 1991; Van Meel, 1991). Auch Köstlin-Gloge (1978) ging im Zusammenhang mit dem Entwicklungsaspekt auf die Dimensionalität von Feldabhängigkeit und Feldunabhängigkeit ein. Der Autorin zufolge handelt es sich bei Feldabhängigkeit zunächst um eine Fähigkeit und erst ungefähr ab dem Alter von 15 Jahren kann von kognitivem Stil im Sinne einer Wahlfreiheit gesprochen werden.

Dem Bestreben, Feldabhängigkeit und Feldunabhängigkeit als gleich zu werten, läuft die Tatsache zuwider, dass Feldabhängigkeit/Feldunabhängigkeit zumeist durch Leistungstests, wie dem EFT, erfasst wird, in denen Feldunabhängigkeit zu besseren Ergebnissen führt und eindeutig als erstrebenswert hervor geht (McKenna, 1984). Feldabhängigkeit ergibt sich hier stets aus einem Mangel an einer bestimmten Fähigkeit. Dass dies mit Stärken in anderen (Fähigkeits)Bereichen einhergeht, wird in der Regel nur behauptet<sup>[s3]</sup>, nicht aber miterfasst. Mittlerweile wird Feldabhängigkeit eher als ein unipolares Fähigkeitskonzept erachtet (Kogan & Block, 1991; McKenna, 1984).

Aufgrund dieser Argumentation soll von nun an der Begriff Feldunabhängigkeit in Bezug auf die entsprechenden Verfahren sowie stellvertretend für die Dimension Feldabhängigkeit/Feldunabhängigkeit verwendet werden. Es steht im Folgenden der Begriff Feldabhängigkeit einzig für einen Mangel an Feldunabhängigkeit.

### 5.3 Freiheit von Gestaltbindung

Thurstone (1938) veröffentlichte eine faktorenanalytische Studie, in der er den zuvor entdeckten *perceptual factor P* bestätigen wollte. Thurstone ging davon aus, dass “the perceptual factor might consist in a facility to perceive detail even when it is buried among perceptual distractors.” Im Zuge dieser Analyse wurde eine Reihe wahrnehmungsbezogener Tests erstellt, wie zum Beispiel *Designs*. In diesem Test muss aus einer Reihe von Linienkombinationen jene gefunden werden, in welcher der Buchstabe  $\Sigma$  eingebettet ist. Die Ähnlichkeit zu den Konzepten der Gottschaldt-Figuren und des EFT ist auffällig. Tatsächlich luden alle speziell erstellten Tests auf dem Faktor *P*, außer *Designs*.

In einer weiteren Faktorenanalyse ging Thurstone (1944) unter anderem auf die Gestalt-Aspekte bestimmter Aufgaben ein. Er sprach von *closure*- oder *gestalt*-Faktoren. Diese Studie war die Basis für weiterführende Arbeiten, welche die Faktoren *speed of closure* und *flexibility of closure* (*Freiheit von Gestaltbindung*) weitgehend bestätigten (s.u.).

Botzum (1951) äußert sich zu seinen Ergebnissen für den Faktor *speed of closure* folgendermaßen: „In all of the tests with loadings on this factor, there is an unstructured field, in which some reorganization must occur. This process of reorganization may well be termed

closure.” (S. 370). So müssen zum Beispiel in *Hidden Pictures* die oft schwach strukturierten Elemente der versteckten Gesichter zu ganzen Gestalten zusammengefügt werden.

Thurstone (1944) nannte den Faktor “*freedom from Gestaltbindung*”, weil “[it] might be an appropriate description in that it implies flexibility in manipulating several more or less irrelevant or conflicting gestalts” (Thurstone, 1944, S. 111). Unter anderem luden die Tests *Blockdesign*, *Designs* und *Hidden Pictures* auf diesem Faktor. In den darauffolgenden Studien gab es kontroverse Ergebnisse bezüglich der Faktorzugehörigkeit von *Hidden Pictures*.

Auch Pemberton (1952) nennt für die Tests auf Faktor  $C_2$  ihrer Studie „*freedom from Gestaltbindung*“ als gemeinsamen Nenner. Sie betont zudem, dass die Zielfigur hierbei stets bekannt ist, anders als beispielsweise bei *Hidden Pictures*, welcher in dieser Studie auf dem Faktor *Speed of Closure* am höchsten lädt. *Freiheit von Gestaltbindung* korreliert dabei mit dem Faktor  $C_3$  *verbal closure*. Pemberton meint hierzu, dass sich im Faktor *verbal closure* Elemente von *Freiheit von Gestaltbindung* finden. Der wesentliche Unterschied scheint in der Vorgabe von verbal-schriftlichem Material zu liegen. *Anagrams* beispielsweise lädt hoch auf *verbal closure* und erreicht auch einen mittleren Wert für *Freiheit von Gestaltbindung*. „In Anagrams the subject must abandon the Gestalt which is formed by one English word and reorganize the letters into another word“ (Pemberton, 1952, S. 275).

Botzum (1951) identifizierte den Faktor  $C$  in seiner Studie mit *Freiheit von Gestaltbindung*. Die *Gottschaldt-Figuren* und *Designs* zeigten hier die 2.- und 3.-höchste Ladung. Dies deckt sich mit den Ergebnissen und der Interpretation von Thurstone (1949). „Tests with high loadings on this factor seem to involve the holding in mind of a configuration or gestalt, and the operating with it against distractors“ (Botzum, 1951, S. 369).

Zusammenfassend kann *Freiheit von Gestaltbindung* also als die Fähigkeit verstanden werden, mit gestaltbildenden Informationen flexibel umzugehen. Darunter fällt entsprechend der besprochenen Studien einerseits die Fähigkeit, gegebene komplexe Gestalten aufzulösen, sowie andererseits die Fähigkeit, die einzelnen Informationselemente eines Bildes verschieden zusammenzusetzen, um letztendlich bekannte Formen rekonstruieren bzw. vergleichen zu können. Klassischer Weise wird dies durch Aufgaben gemessen, bei denen eine vorgegebene Figur in einem Bild wiederentdeckt werden muss, in dem eine Vielzahl konkurrierender Informationen und entgegenlaufender Gestalt-Prinzipien das Herausfiltern der relevanten Figur erschweren.

## **Freiheit von Gestaltbindung und Feldunabhängigkeit**

Die Definition von *Freiheit von Gestaltbindung* sowie die Aufgabenstellung der entsprechenden Tests – insbesondere *Designs* – weisen auffällige Parallelen mit jenen der *Gottschaldt-Figuren* und des EFT auf.

Wie bereits erwähnt, fanden Goodenough und Karp (1961) einen Zusammenhang zwischen Feldunabhängigkeit und dem *Blockdesign*-Test, welcher auf *Freiheit von Gestaltbindung* lädt. Auch andere Studien konnten derartige Verbindungen zu den auf *Freiheit von Gestaltbindung* ladenden Verfahren finden (Gardner, Jackson und Messick, 1960; Young, 1959). Bei Gardner et al. (1960) zeigte sich zum Beispiel, dass Tests, die mit *Freiheit von Gestaltbindung* identifiziert werden (*Concealed Figures* und *Designs*), und Tests, die für Feldunabhängigkeit stehen (EFT und RFT), auf demselben Faktor laden.

“Taken together, these studies provide impressive support for the view that flexibility of closure, spatial decontextualization, and field dependence may be different names for the same dimension” (Witkin et al., 1962, S. 52).

Witkins Schlussfolgerung kann hier – mit Ausnahme der Dimensionsbezeichnung – nur zugestimmt werden.

## **6. Schlussfolgernde Zusammenfassung**

Es stellt sich heraus, dass Feldunabhängigkeit und *Freiheit von Gestaltbindung* nichts anderes als verschiedene Bezeichnungen für ein und dasselbe Konstrukt sind. Deren operante Definitionen decken sich beide fast vollständig mit der des UT 10 und wie sich bei Hergovich (1992) zeigte, muss von einer Gemeinsamkeit mit dem in UT 10 erfassten Konstrukt ausgegangen werden.

Einige experimentelle Studien konnten einen wichtigen Beitrag zum Verständnis der Bedeutung von bestimmten Mustereigenschaften für die mentale Verarbeitung der Aufgaben erbringen. Die Ergebnisse stützen die Annahmen über die Konstruktvalidität des UT 10. Das Prinzip der Prägnanz spiegelte sich in vielen Beschreibungen der Muster wider. Die Theorien der Gestaltpsychologie erlauben Annahmen über die Beziehung bestimmter Mustereigenschaften und deren Wahrnehmung als Figur bzw. Gestalt. Sie geben Hinweise darauf, welche Reizeigenschaften bei bestimmten Arten von Abbildungen von Bedeutung

sind. In der *Structural Information Theory* bietet sich eine Möglichkeit, diese Beziehung bei simplen Formen zu erfassen.

Aus der Auseinandersetzung mit den genannten Studien lässt sich ableiten, dass es sich bei der analytischen Vorgehensweise beim Bearbeiten von *Blockdesign*-Aufgaben um den Ausdruck einer höherrangigen Fähigkeit handelt, welche der gewünschten *Strukturierung* im UT 10 entspricht. Mit höherrangig ist dabei einerseits gemeint, dass sie der globalen Vorgehensweise in bedeutenden Aspekten (v.a. Bearbeitungsgeschwindigkeit) überlegen ist, und andererseits, dass es den Personen, welche global vorgehen, scheinbar nicht *möglich* ist, die Aufgaben analytisch zu lösen. Während also Personen, welche über ein gewisses Maß dieser Fähigkeit verfügen, ihre Vorgehensweise den Gegebenheiten anpassen und zwischen globaler und analytischer Bearbeitung wechseln können, sind Personen mit einer nur sehr geringen Ausprägung dieser Fähigkeit auf die globale Vorgehensweise beschränkt. Ein gewisser Entwicklungsaspekt dieser Fähigkeit kann ebenfalls identifiziert werden.

Die Wahl(möglichkeit) der angewandten Strategie sowie folglich der Erfolg der Itembearbeitung scheint mit der Ausprägung von Feldunabhängigkeit einher zu gehen. Laut Hergovich (1992) liegt dieses Konstrukt dem UT 10 zugrunde.

Der besondere Vorteil einer vorwiegend analytischen Vorgehensweise liegt, wie sich zeigte, im geringeren Zeitaufwand. Dies hat eine bedeutende Auswirkung für die Lösung der Aufgaben. Die Strategie erhält durch die Zeitbeschränkung der Aufgaben einen gewissen Notwendigkeitsstatus. Denn selbst wenn durch die globale Strategie das Muster nachgelegt werden kann, führt die Zeitbeschränkung womöglich zu einer Wertung als „nicht gelöst“. Sowohl durch diese Zeitkomponente, als auch durch die Möglichkeit, die Aufgaben durch verschiedene Strategien gleichermaßen lösen zu können, ergeben sich mögliche Implikationen bezüglich des für den UT 10 geltenden *Rasch*-Modells, welche in Kapitel 12 diskutiert werden.

Was sich im Überblick ebenfalls zeigt ist, dass eine gewisse Beteiligung von mentaler Rotation für den Lösungsprozess angenommen werden muss. Zweifarbig Würfel stellen insofern eine höhere Herausforderung dar, als einfarbige. Der Begriff der *guten Gestalt* wurde mehrfach mit dem Zergliederungsprozess in Verbindung gebracht und stellt demnach eine mögliche Schwierigkeitsdeterminante für Aufgaben der Sorte *Blockdesign* dar.

## IV. Testtheoretischer Rahmen des linear logistischen Testmodells

---

Das *Adaptive Intelligenz Diagnostikum 3* orientiert sich an der modernen Testtheorie (auch probabilistische Testtheorie oder Item-Response Theory, kurz IRT). Basis für das *branched testing*, welches im UT 10 angewandt wird, stellt die Geltung des *Rasch-Modells* (Rasch, 1980) dar. Gleichzeitig ist dies die Voraussetzung für die Anwendung des Linear Logistischen Testmodells von Fischer (1973), welches hier angewendet werden soll. Daher werden hier sowohl die Grundlagen und Bedingungen als auch die Möglichkeiten zur Prüfung dieser beiden Modelle genauer dargestellt.

### 7. Das dichotom logistische Testmodell von Rasch (*RM*)

Den Latent-Trait-Modellen der probabilistischen Testtheorie liegt die Idee zugrunde, dass die Wahrscheinlichkeit einer bestimmten Antwort einer Person auf ein Item idealer Weise nur von der Ausprägung der zu messenden latenten Eigenschaft (*latent trait*) der Person sowie einem oder mehreren Parametern, welche das Item charakterisieren, abhängt. Der Zusammenhang zwischen der Ausprägung der latenten Eigenschaft  $\xi$  und der Wahrscheinlichkeit einer bestimmten Antwort wird als systematisch angenommen und für jedes einzelne Item  $i$  durch die Itemcharakteristik  $f_i(\xi)$  beschrieben. Diese besagt, dass für alle Personen mit derselben Ausprägung  $\xi$  auch dieselbe Wahrscheinlichkeit für diese Antwort besteht (Fischer, 1974).

Bei Latent-Trait-Modellen wird generell *lokale stochastische Unabhängigkeit* angenommen. Außerdem erfassen alle Items eines Verfahrens nur eine einzige latente Dimension (Eigenschaft). Die Beantwortung eines Items ist allein von dieser einen latenten Dimension abhängig, aber unabhängig davon, wie viele und welche anderen Items bereits bearbeitet bzw. gelöst wurden. Lediglich die Abhängigkeit von einer gemeinsamen latenten Dimension determiniert die Kovariation zwischen den Antworten auf einzelne Items (Fischer, 1974). Unter gewisser Voraussetzung kann ein eindimensionales Latent-Trait-Modell auch für Items mit zwei oder mehreren zugrundeliegenden latenten Dimensionen Gültigkeit finden; nämlich

dann, wenn alle Items dieselbe Kombination dieser Dimensionen erfassen (Reckase, 1979, zitiert nach Embretson & Reise, 2000).

Das wohl bekannteste und am häufigsten angewandte eindimensionale Latent-Trait-Modell ist das dichotom logistische Testmodell von Georg Rasch (1980) (zumeist und auch hier der Einfachheit halber im Weiteren als *Rasch*-Modell oder *RM* bezeichnet). Dichotom bedeutet, dass die Antwort auf ein Item nur negativ („0“) oder positiv („1“) ausgeprägt sein kann. In der Regel wird davon gesprochen, ob eine Aufgabe von einer Person „nicht gelöst“ oder „gelöst“ wurde. Diese Formulierung lässt vermuten, dass das *Rasch*-Modell sich allein auf Leistungstests bezieht. Wenngleich dies das Hauptanwendungsgebiet darstellt, ist dem nicht so (Fischer, 1974). Weil die vorliegende Arbeit sich aber tatsächlich mit einem Leistungstest auseinandersetzt, werden für die weitere Abhandlung die Begriffe „Aufgabe“ mit „Item“ und „latente Eigenschaft“ mit „Fähigkeit“ ident gesetzt. Die folgende Darstellung beziehen sich, sofern nicht anders angegeben, auf Fischer (1974).

Das *Rasch*-Modell kann mit folgender Formel zusammengefasst werden<sup>2</sup>:

$$P(+|\xi_v; \sigma_i) = \frac{e^{\xi_v - \sigma_i}}{1 + e^{\xi_v - \sigma_i}} \quad (1)$$

Die Wahrscheinlichkeit  $P$ , dass die Person  $v$  Item  $i$  löst („+“), ist demnach ausschließlich von der Differenz zwischen der (wahren) Fähigkeit von Person  $v$ , dargestellt als Personenparameter  $\xi_v$ , und der (wahren) Schwierigkeit von Item  $i$ , dargestellt als Itemparameter  $\sigma_i$ , abhängig. Daraus lässt sich schließen, dass bei Gültigkeit des *Rasch*-Modells von der Eindimensionalität der Items ausgegangen werden kann (Kubinger, 2009b).

Ein besonderer Vorteil des *Rasch*-Modells ist die sogenannte *erschöpfende Statistik*: Die Summe der positiven Antworten einer Person liefert ausreichend Information über den Personenparameter und die Summe der positiven Antworten auf ein Item liefert ausreichend Information über den Itemparameter. Dies vorausgesetzt ergeben sich weitere bedeutende Vorteile wie Stichprobenunabhängigkeit (auch Populationsunabhängigkeit) und die *spezifische Objektivität*. Nämlich ist die Schätzung der Parameter unabhängig von der

---

<sup>2</sup> Darstellung mit logarithmierten Parametern

Verteilung der wahren Personenparameter in der benützten Stichprobe. Die Stichprobe muss also nicht repräsentativ sein.<sup>3</sup> Findet man bedeutende Unterschiede, so muss das Modell für die Gesamtpopulation verworfen werden. Gilt das *Rasch*-Modell, sind also spezifisch objektive Vergleiche möglich, was es ermöglicht, den Fähigkeitsunterschied zweier Personen unabhängig davon, welche Aufgaben bearbeitet wurden, zu bestimmen, sowie die Schwierigkeiten zweier Aufgaben unabhängig davon, welche Stichprobe verwendet wird, zu vergleichen (Kubinger, 2009). Dies bildet die Grundlage des adaptiven Testens und gewährleistet eine faire Verrechnung.

## 8. Die Prüfung der Geltung des RM

Um die  $H_0$  „Das Modell gilt“ zu prüfen wird die Gesamtstichprobe anhand von Teilungskriterien in (üblicher Weise) zwei Gruppen geteilt und je Teilstichprobe eine Schätzung der Itemparameter durchgeführt. Wenn die Itemparameter der beiden Teilstichproben und der Gesamtstichprobe gleich sind ( $\sigma_{i1} = \sigma_{i2} = \sigma_i$ ), kann die Geltung der  $H_0$  angenommen werden (Fischer, 1974). Als Teilungskriterien können sowohl externe Kriterien, wie Geschlecht, Alter, Schultyp oder Muttersprache, als auch das interne Kriterium Score (hoch versus niedrig) dienen. Üblicher Weise wird der Datensatz anhand von 3 bis 5 Teilungskriterien untersucht, wobei Score, Geschlecht und Alter meist zur Auswahl gehören (Kubinger, 2005).

In der vorliegenden Arbeit werden der *Likelihood Ratio Test* nach Andersen (1973) sowie graphische Kontrollen zur Prüfung der Modelle angewandt.

### 8.1 Der Likelihood Ratio Test nach Andersen beim RM

Für die Schätzung der Parameter wurden verschiedene Methoden entwickelt, die von Computerprogrammen berechnet werden können. Standardmäßig wird zur Prüfung der Geltung des *Rasch*-Modells der *Likelihood Ratio Test (LRT)* nach (Andersen, 1973; auch Andersen *Likelihood-Quotienten Test* genannt, kurz *LQT*) herangezogen, der alle Items simultan einbezieht (Kubinger, 2005). So auch bei der Prüfung des Untertests *Analysieren*

---

<sup>3</sup> Stichprobenunabhängigkeit bedeutet nicht, dass die gewählte Stichprobe keinerlei Einfluss auf das Ergebnis hat. Die statistische Genauigkeit, mit der die Parameter geschätzt werden, hängt sehr wohl von der Stichprobengröße und einer angemessenen Verteilung von Lösungshäufigkeiten je Item ab (Fischer, 1974).

und Synthesieren durch Kubinger und Holocher-Ertl (in Druck). Der *LRT* basiert auf dem Konzept der *Conditional Maximum Likelihood* (*CML*). Einfach formuliert ist hier die *Likelihood* „das Produkt der Wahrscheinlichkeiten für die beobachteten Daten bei gegebenen Parametern“ (Koller, Alexandrowicz & Hatzinger, 2012, S. 67).

Bei Gültigkeit des *Rasch*-Modells ergibt sich, dass

$$2 * (\ln L_1 + \ln L_2 - \ln L_g) \quad (2)$$

$L_g$  .... Likelihood der Daten der Gesamtstichprobe

$L_1$  .... Likelihood der Daten der Teilstichprobe 1

$L_2$  .... Likelihood der Daten der Teilstichprobe 2

bei  $df = k_1 + k_2 - k_g$

$k_g$  .... Anzahl der geschätzten Itemparameter in der Gesamtstichprobe

$k_1$  .... Anzahl der geschätzten Itemparameter in der Teilstichprobe 1

$k_2$  .... Anzahl der geschätzten Itemparameter in der Teilstichprobe 2

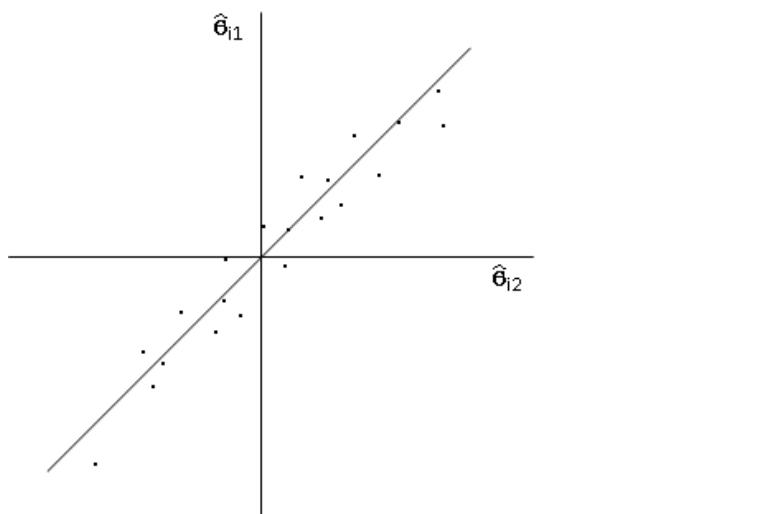
mit genügend großer Stichprobe asymptotisch  $\chi^2$ -verteilt ist. Ist der Ergebniswert also kleiner als der kritische  $\chi^2$ -Wert, kann für das getestete Teilungskriterium die  $H_0$  beibehalten werden (Andersen, 1973). Der *LRT* stellt eine sehr genaue Methode für die Modellkontrolle dar. Allerdings werden insbesondere bei großen Stichproben sehr häufig Modellabweichungen angezeigt (Näherer, 1977, 1980; Kubinger, 1979).

## 8.2 Die graphische Modellkontrolle beim *RM*

Eine weitere, anschauliche, Möglichkeit stellt die graphische Modellkontrolle (Abbildung 4) dar, bei der in einem rechtwinkligen Koordinatensystem die Parameterschätzungen beider Teilstichproben je Item mit einem Punkt  $P(\hat{\theta}_{i1}, \hat{\theta}_{i2})$  repräsentiert werden (Formann, 2008). Bei Gültigkeit des *Rasch*-Modells liegen (im theoretischen Idealfall) alle Punkte auf der  $45^\circ$ -Geraden.

Die graphische Kontrolle ist insbesondere dann hilfreich, wenn die Ergebnisse des globalen inferenzstatistischen Tests auf eine Modellverletzung hinweisen. Es lässt sich leicht erkennen, ob sich der Punktschwarm generell nicht entlang der  $45^\circ$ -Geraden verteilt, oder ob nur einzelne Items stärker abweichen. In letzterem Falle kann unter Umständen durch das Ausscheiden eines oder mehrerer einzelner Items doch noch Modellgültigkeit erreicht werden [s4].

Welche Abweichungen von der erwarteten Geraden signifikant sind, ist dabei in dieser einfachen Darstellung nicht zu erkennen.



**Abbildung 4:** Beispiel für die graphische Kontrolle des Rasch-Modells. Die Itemparameterschätzungen  $\hat{\theta}_{i1}$  und  $\hat{\theta}_{i2}$  der beiden Teilstichproben 1 und 2 stimmen für alle  $i$  Items fast genau überein.

## 9. Das linear logistische Testmodell (*LLTM*) von Fischer

Das linear logistische Testmodell von Fischer (1973) (im Weiteren mit der Abkürzung *LLTM* bezeichnet) kann als eine Erweiterung des *Rasch*-Modells verstanden werden und ermöglicht es, die Differenz zwischen den Itemparametern zu erklären. Es wird also erfasst, was für die Schwierigkeit der Aufgaben verantwortlich ist bzw. was ein Item schwieriger oder leichter macht [s5] als ein anderes. Die folgende Beschreibung des Modells bezieht sich wieder, sofern nicht anders angegeben, auf Fischer (1974).

Angenommen wird einerseits, dass sich die Itemschwierigkeit aus der Beanspruchung bestimmter kognitiver Operationen ergibt (Schreiblechner, 1972 zitiert nach Fischer, 1995), und andererseits, dass Bedingungen der Testsituation bzw. formale Itemeigenschaften (z.B. thematischer Inhalt oder Farb- und Strukturmerkmale) eine Auswirkung auf die Itemschwierigkeit haben. Im *LLTM* wird der Schwierigkeitsparameter  $\sigma_i$  in eine gewichtete Summe von *Basisparametern*  $\eta_j$  zerlegt, welche den Schwierigkeiten dieser hypothetisierten kognitiven Operationen bzw. Bedingungen entsprechen.

Während in einschlägiger Literatur weithin der Begriff Operation oder *Elementaroperationen* (z.B. Fischer & Formann, 1982, Kubinger, 2008) verwendet wird, scheint dies unter Berücksichtigung von Testbedingungen und formalen Itemeigenschaften eher irreleitend. Der Begriff *Operation* soll im Folgenden nur mehr für kognitive Prozesse verwendet werden. Zusätzlich wird der Begriff *Elementarfaktor* eingeführt, der für die auf das *LLTM* bezogenen Testbedingungen, Itemeigenschaften und Operationen zusammenfassend verwendet wird.

In einer Formel ausgedrückt ist das *LLTM* ein *Rasch*-Modell mit folgender Nebenbedingung:

$$\sigma_i^* = \sum_j^m q_{ij} \eta_j + c \quad (3)$$

$\sigma_i$  ..... Schwierigkeit von Item  $i$

$\eta_j$ ..... Schwierigkeit von Elementarfaktor  $j$

$q_{ij}$ .... Gewichtszahl des  $j$ -ten Elementarfaktors bei Item  $i$

$c$  ..... Normalisierungskonstante

Die Schwierigkeit von Item  $i$  ergibt sich also aus der Summe der für das Item gewichteten Schwierigkeiten je Elementarfaktor. Die Normalisierungskonstante wird in der Regel aus der Formel genommen, da die Schätzung der Parameter davon unabhängig ist. In das *Rasch*-Modell integriert ergibt sich damit folgende Formel für das *LLTM*:

$$P(+|\xi_v; \sigma_i = \sum_j^m q_{ij} \eta_j) = \frac{e^{\xi_v - \sum_j^m q_{ij} \eta_j}}{1 + e^{\xi_v - \sum_j^m q_{ij} \eta_j}} \quad (4)$$

Hieraus lässt sich auch ablesen, dass die Voraussetzung für die Gültigkeit des *LLTM* die Gültigkeit des *Rasch*-Modells ist. Zusammengefasst besagt das *LLTM*, dass Unterschiede zwischen Itemparametern sich aus Unterschieden in der Anwendung bestimmter kognitiver Operationen (oder in Effekten gewisser Bedingungen) zwischen den Items ergeben.

Welche Eigenschaften der Items nun aber überhaupt eine Rolle spielen muss, wie erwähnt, zunächst einmal in einer Hypothese erfasst und dann operational definiert werden. Idealer Weise wird bereits bei der Entwicklung eines Verfahrens mitbedacht oder festgelegt, welche logischen Operationen zur Lösung führen bzw. welche Aufgabenbedingungen deren Beantwortung beeinflussen könnten. Ist dies nicht der Fall, wird das bestehende Material einer eingehenden Analyse unterzogen. Können über die Beteiligung bestimmter mentaler Prozesse Hypothesen aufgestellt werden, gilt es, diese für deren Prüfung in messbare und angemessen gewichtete (skalierte) Elementarfaktoren umzuformulieren. Die Bewertungen aller Items anhand dieser Elementarfaktoren werden in eine Strukturmatrix eingetragen, die bestimmte Eigenschaften zu erfüllen hat.

## 9.1 Die Strukturmatrix des *LLTM* (*Q*-Matrix)

Der Kern der *Strukturmatrix* (auch *Gewichtszahlenmatrix* oder *Q*-Matrix) besteht aus  $k$  Zeilen und  $m$  Spalten, wobei  $m$  die Anzahl an Basisparametern, also Elementarfaktoren, repräsentiert und  $k$  die Anzahl an Items (Tabelle 1). Hierbei gilt stets  $m < k$ . Es sollen die Itemparameter durch möglichst wenige, nämlich nur die relevanten, Elementarfaktoren vorhergesagt werden können.

In die Felder werden nun je Item und Elementarfaktor die hypothetisch festgelegten Gewichtszahlen eingetragen. Zumeist ergeben sich die Gewichtszahlen aus einer dichotomen Kodierung. Muss die jeweilige Operation bei einem Item angewendet werden bzw. hat dieses Item die jeweilige Eigenschaft wird „1“ eingetragen, wenn nicht „0“. Muss eine Operation bei einem Item mehrmals angewendet werden, so geht die Häufigkeit als Gewichtung in die Matrix ein. Eine *Q*-Matrix kann also aus mehr als nur Nullen und Einsen bestehen. Zu berücksichtigen ist jedoch, dass die Spaltenvektoren voneinander linear unabhängig sein müssen, also vollen Rang besitzen. In der Matrix sollten dadurch keine überflüssigen Parameter enthalten sein (Fischer, 1974).

**Tabelle 1:** Schematische Darstellung der Gewichtszahlenmatrix des *LLTM*

		kognitive Operationen				
		1	...	$j$	...	$m$
Item	1					
	$i$	$q_{i1}$		$q_{ij}$		$q_{im}$
	$k$					

$\sigma_i$

## 9.2 Zur Anwendung des *LLTM*

Die Festlegung aller relevanter Operationen und Einflüsse ist nie einfach und gelingt nur selten. Auch nur geringe Missspezifizierungen in der Matrix können aber bereits große Auswirkungen auf die Schätzwerte haben (Baker, 1993). Gittler und Wild (1989) fassen daher einige wichtige Aspekte für den Einsatz des *LLTM* zusammen: Bereits zu Beginn ist zu überlegen, ob eine lineare Zerlegung der Itemparameter überhaupt Sinn hat. Nicht alle Itemmaterialien sind dafür gleichermaßen geeignet. Außerdem bedarf es präziser Hypothesen über die Elementarfaktoren und deren Beziehung sowohl zum Itemmaterial als auch untereinander. Deren Definitionen sollten gründliche, inhaltspsychologische Analysen zugrunde liegen. Sowohl theoretische und praktische Auseinandersetzung mit dem Material als auch genaue Beobachtungen der Aufgabenbearbeitung sollten dem Aufstellen der Hypothesen vorangehen. Als das Wesentliche bezeichnen Gittler und Wild dabei, dass die Schwierigkeitsunterschiede zwischen den Aufgaben möglichst genau durch die Elementarfaktoren abgebildet werden. Das bedeutet trivialer Weise, dass nur diejenigen einbezogen werden, welche an den Aufgaben verschieden stark beteiligt sind und erst damit zur Differenzierung beitragen können.

Im Zusammenhang damit steht auch die Qualität des Itempools. Damit itemgruppenspezifische Parameterschätzungen vermieden werden können, sollte ein möglichst großer und bezüglich individueller Eigenschaftszusammensetzungen ausreichend heterogener Itempool zur Verfügung stehen, sodass die einzelnen Elementarfaktoren

möglichst vollständig kombiniert darin vorkommen (Hornke & Rettig, 1989). Das heißt, eine gemeinsame Variation zweier oder mehrerer Charakteristika ist als problematisch anzusehen.

Aus der Möglichkeit, die für die Lösung der Aufgaben relevanten Elementarfaktoren empirisch zu prüfen, ergeben sich verschiedene Anwendungsgebiete für das *LLTM*. Zwei bedeutende Einsatzmöglichkeiten sind die Konstruktvalidierung sowie die Generation von Itemkonstruktionsregeln (Kubinger, 2008).

### **Konstruktvalidierung mittels *LLTM***

Entsprechend der Idee, dass ein Test dann per Definition valide ist, wenn zur Lösung der Aufgaben genau jene Kompetenzen erforderlich sind, welche durch den Test erfasst werden sollen, lässt sich mit dem *LLTM* die Konstruktvalidität eines Tests überprüfen. Wenn bestimmte kognitive Operationen angewendet werden müssen und die Lösung auf keine andere Art erzielt werden kann, dann kann von einem hohen Testscore auf eine hohe Kompetenz in der Durchführung dieser kognitiver Operationen geschlossen werden (Kubinger, 2008).

Zumeist jedoch zeigt sich, dass auch formale Bedingungen und Itemeigenschaften mit für die Schwierigkeit der Aufgaben verantwortlich sind, so dass ein gültiges *LLTM* in der Regel nur unter Einbezug solcher formalen Kriterien gelingt. Sonnleitner (2008) zum Beispiel überprüfte mittels *LLTM* kognitive Modelle zum Leseverständnis. Im Rahmen dessen stellte sich auch die Anzahl möglicher Antworten im Multiple Choice-Format als relevanter Elementarfaktor für die Testung von Leseverständnis heraus. Auch Poinstingl (2008) konnte für das *LLTM* des untersuchten *Reasoning*-Tests trotz mangelnder Modellgültigkeit bessere Ergebnisse unter Einbezug formaler Elementarfaktoren erzielen und damit auch zumindest einen richtungsweisenden Beitrag zur Generierung valider Itemkonstruktionsregeln für den Test leisten.

### **Regelgeleitete Itemkonstruktion mittels *LLTM***

Durch die Prüfung der relevanten Itemeigenschaften ist es theoretisch innerhalb des festgelegten Geltungsbereiches möglich, beliebig komplexe und viele Items zu konstruieren sowie deren Schwierigkeit vorherzusagen. Im Grunde können damit auch gezielt für Schwierigkeitsgrade, für die noch nicht genügend Items vorliegen, zusätzliche Items erstellt

werden. Einer Testperson eine entsprechend ihrer vorläufig geschätzten Fähigkeit schwierige Aufgabe vorlegen zu können, ist Voraussetzung für adaptives Testen. Beim computerisierten Testen ist es mit den passenden Regeln theoretisch möglich, während des Testens für die laufende Schätzung maßgeschneiderte Items zu konstruieren und somit den maximalen Nutzen des adaptiven Testens zu erreichen (Kubinger, 2008). Das Erstellen von Paralleltests wird durch die Möglichkeit, gezielt gleichwertige Items produzieren zu können, ebenfalls wesentlich erleichtert.

## 10. Die Prüfung der Geltung des *LLTM*

Für die Prüfung des *LLTM* werden weitgehend dieselben Methoden angewendet, wie bei der Prüfung des *Rasch-Modells*. Als Vergleichsgrundlage dienen nun aber die Itemparameter des *RM* und des *LLTM* statt der jeweiligen Teilstichproben.

Wenn die festgelegten Elementarfaktoren der *Q-Matrix* tatsächlich für die Schwierigkeit der einzelnen Items verantwortlich sind, und zwar in dem Maße, wie sie durch die Gewichtung bestimmt werden, dann lässt sich durch die *Q-Matrix* die Schwierigkeit der einzelnen Items vorhersagen. Die Schätzung der Itemparameter durch das *LLTM* sollte also der des *RM* entsprechen. Geprüft wird damit nicht, ob das *LLTM* an sich auf die Daten passt, sondern ob die Nullhypothese der Nebenbedingungen aufrechterhalten werden kann (Fischer, 1974).

### 10.1 Der *Likelihood Ratio Test* nach Andersen beim *LLTM*

Für die inferenzstatistische Prüfung wird wiederum der kritische  $\chi^2$ -Wert herangezogen und mit dem Ergebniswert des *LRT* verglichen. Es gilt:

$$-2 \ln (L_{LLTM} / L_{RM}) \quad (7)$$

bei  $df = k - m - 1$

$k$  ..... Anzahl der geschätzten Itemparameter im *RM*

$m$  ..... Anzahl der Basisparameter im *LLTM*

ist bei einer ausreichend großen Stichprobe asymptotisch  $\chi^2$ -verteilt (Fischer, 1974).

## 10.2 Die graphische Modellkontrolle beim *LLTM*

Hierfür gilt dasselbe Prinzip wie für die graphische Modellkontrolle des *Rasch*-Modells, nur werden nicht die Itemparameterschätzungen je Teilstichprobe  $P(\hat{\theta}_{i1}, \hat{\theta}_{i2})$  im Koordinatensystem e getragen, sondern die Itemparameterschätzungen der beiden Modelle  $P(\hat{\theta}_{iRM}, \hat{\theta}_{iLLTM})$  gepaart. Es gilt wieder: Je näher die sich ergebenden Punkte an der  $45^\circ$ -Geraden liegen, desto besser. Dies zeigt nämlich, wie gut das *LLTM* die Itemparameter des *RM* vorhersagen kann.

## V. Modellerstellung

---

### 11. Vorgehensweise

Vor der intensiven Beschäftigung mit der Konstruktvalidität des UT 10 und den Untersuchungen über die kognitiven Prozesse bei dessen Bearbeitung, wurde sich dem Material erst einmal „naiv“ gewidmet. Die Verfasserin unterzog sich der Vorgabe in der Position als Testperson und verfasste dabei introspektive Beobachtungen und erste Annahmen über die möglichen kognitiven Operationen und relevanten Materialeigenschaften in Form von Notizen, wobei sich auch genauer mit den Mustereigenschaften befasst wurde. Die strukturierte Zusammenfassung dieser *Testperson-Notizen (TPN)* ist in Tabelle 18 im Anhang wiedergegeben. Selbstverständlich muss dabei ein anderes Denken und Erleben als bei Kindern- und Jugendlichen in einer Testsituation angenommen werden. Dennoch wird hier davon ausgegangen, dass dieses Vorgehen einen Einblick in die mentalen Prozesse bei der Bearbeitung der Aufgaben gewähren und die Gefahr, relevante Aspekte zu übersehen, reduzieren kann.

Die Verfasserin beteiligte sich zudem an Erhebungen für die Äquivalenzprüfung der Gruppenversion des AID 3 (Kubinger & Hagenmüller, in Arbeit), welche vom Arbeitsbereich Psychologische Diagnostik der Universität Wien, durchgeführt wurden. Neben Einsichten durch die Vorgabe der Untertests 10 und 10a im Einzelsetting konnten auch Erfahrungen durch das Beiwöhnen und Unterstützen bei einer Gruppentestung gesammelt werden. Bei Gesprächen mit Kollegen wurden unter anderem Eindrücke und Gelegenheitsbeobachtungen zur Bearbeitung und Vorgabe der dieser beiden Untertests ausgetauscht. Aus den Beobachtungen wurden gemeinsam mit den Inhalten dieses Austausches weitere Annahmen über den Lösungsprozess entwickelt. Die strukturierte Zusammenfassung der *Testleiter-Notizen (TLN)* hiervon ist im Anhang in Tabelle 19 vollständig enthalten.

Es handelt sich bei diesen Notizen zwar nicht um streng systematisch gesammelte Informationen, doch entspricht durch diese Art der Herangehensweise die Arbeit den Forderungen von Gittler und Wild (1989, in Kapitel 9.2 dargestellt), wonach dem Aufstellen der Hypothesen neben der theoretischen Auseinandersetzung mit dem Material auch eine

praktische Beschäftigung damit sowie genaue Beobachtungen der Aufgabenbearbeitung vorangehen sollen. Zudem zeigen sich rückblickend erstaunlich große Übereinstimmungen mit den in Abschnitt III berichteten Untersuchungen, sowohl was die Identifikation von Bearbeitungsstrategien (*TPN* 13 und 14) anbelangt, als auch was die vermutete Relevanz gewisser Materialeigenschaften (*TPN* 11, 12, 15 und 16) betrifft. Die Gelegenheitsbeobachtungen bei der Vorgabe des UT *10a* (*TLN* 6 und 8) hingegen liefern wertvolle zusätzliche Informationen über die Strukturierungsleistung. Daher sollen die *Testperson-Notizen* und die *Testleiter-Notizen* zur Hypothesengenerierung ergänzend mitberücksichtigt werden. Kern der Hypothesengenerierung bilden die aus den Abschnitten II und III hervorgehenden Kenntnisse über mentale Gliederung.

## 12. Modellrahmen und Hypothesengenerierung

Es wurden in Abschnitt II mehrere Untersuchungen besprochen, welche die Existenz von mindestens zwei bedeutenden Strategien bei der Bearbeitung von *Blockdesign*-Aufgaben und ähnlichen Tests nahelegen. Dass die Muster des UT *10* anscheinend auf verschiedene Arten und durch unterschiedliche mentale Prozesse erfolgreich nachgelegt werden können, widerspricht an sich Kubingers (2008) Forderung, dass die Lösung nur auf eine bestimmte Weise erzielt werden können soll. Sowohl dies als auch die hinzukommende Zeitkomponente durch die Zeitlimits bei der Vorgabe des Tests spricht auf den ersten Blick für eine mögliche Verletzung der Eindimensionalität. Wie sich allerdings zeigte, liegt der bedeutendste Vorteil einer stärker analytischen Vorgehensweise im geringeren Zeitaufwand für die Bearbeitung. Außerdem ist bereits bekannt, dass für den Test Eindimensionalität angenommen werden kann (Kubinger & Holocher-Ertl, in Druck). Betrachtet man das alles zusammen, so ergibt sich folgende mögliche Erklärung, für die *Rasch*-Modell-Konformität des UT *10*:

Die Eindimensionalität kann nicht trotz der Zeitbegrenzung angenommen werden, sondern wird erst durch sie gesichert. Wenn ein rein globales Vorgehen ebenfalls zur korrekten Reproduktion der Muster führen kann, dafür aber bedeutend mehr Zeit benötigt wird, dann kann davon ausgegangen werden, dass durch die Limitierung der zur Verfügung stehenden Zeit ein solch reines globales Vorgehen nicht zur positiven Wertung führt. Alleine Personen, die analytisch vorgehen, schaffen es, das Item zeitgerecht zu lösen. Nur durch ein gewisses Mindestmaß an Strukturierung können die Muster zeitgerecht nachgelegt werden. Somit kann wieder angenommen werden, dass die Lösung der Items nur auf eine Weise, nämlich die analytische, möglich ist. Dies entspricht auch der operanten Definition des Untertests, wonach

eben nicht die Fähigkeit gemessen wird, die Muster reproduzieren zu können, sondern die Fähigkeit, die Muster „durch eine geeignete Strukturierung reproduzieren zu können.“ (Kubinger, 2009a, S. 11).

Als Kern der analytisch-strukturierenden Vorgehensweise kristallisiert sich die Bildung mentaler Raster heraus, im Gegensatz zur globalen Bearbeitung. Sowohl je Würfel als auch für das Muster im Gesamten werden demnach mental die Linien der Würfelkanten auf das Muster gelegt. Interessante Beobachtungen hierzu sind aus der Vorgabe des UT *10a* bekannt (*TLN 6 und 8*). Demnach ziehen Kinder, die sich sowohl im UT *10* als auch im UT *10a* als sehr erfolgreich herausstellten, im UT *10a* Raster für das gesamte Muster, indem sie durchgehende Linien ziehen. Andere, ebenfalls erfolgreiche, Kinder zeichnen die Ränder der Würfelkanten einzeln für jeden Würfel im Muster nach, oder gehen kombiniert vor, indem sie für manche Würfel einzeln die Ränder einzeichnen und daran ausgerichtet weitere Linien ziehen. Es zeigte sich in der Auseinandersetzung mit dem Thema, dass gewisse Bildinformationen den Prozess des Rasterziehens erleichtern und andere hingegen dessen Aufrechterhaltung stören können. Zusammenfassend scheinen folgende zwei Arten an Bildinformationen für die Strukturierungsleistung, also das Erkennen der Ränder einzelner Würfelseiten, ausschlaggebend zu sein:

- das Maß an Übereinstimmung der Musterstruktur mit der faktischen Struktur der Würfelkombinationen
- die Stärke der inhärenten Struktur des Musters im Sinne einer als Einheit wahrgenommenen Fläche

Zusätzlich scheint die Art der verwendeten Würfelseiten eine weitere kognitive Operation einzufordern. Nämlich kann für die zweifarbigen Oberflächenelemente angenommen werden, dass diese nur durch mentale Rotation als zur Verfügung stehende Alternative identifiziert werden können (Schorr et al., 1982; Rozencwajg & Corroyer, 2001).

Insgesamt wird damit davon ausgegangen, dass der Untertest *Analysieren und Synthetisieren* erstens die abstrakte Gliederung vorstrukturierter Muster fordert sowie zweitens die zweidimensionale Rotation von Bildelementen anhand von zwei Achsen. Die Gliederung wird dabei als Prozess gesehen, der sich aus mehreren mentalen Operationen zusammensetzt.

## 13. Hypothesen und Elementarfaktoren

Bevor ein Überblick über die einzelnen Elementarfaktoren gegeben wird, soll die Relevanz des besprochenen theoretischen Rahmens für die Operationalisierung der einzelnen Hypothesen angesprochen werden. Leider zeigte sich, dass eine Übertragung der theoretischen Annahmen und Methoden auf das konkrete Itemmaterial nicht in gewünschtem Maße möglich war.

Für die Wahrnehmung der Struktur des Vorlagenmusters ist entsprechend der gesichteten Literatur im Besonderen das gestaltpsychologische Metaprinzip der Prägnanz von Bedeutung. Im Sinne von Symmetrie und Regelmäßigkeit sowie Gestalt- bzw. Figurcharakter sollten die einzelnen Figur-Grund-Relationen sowie Symmetrie zur Beschreibung der Vorlagemuster dienen. Dieses Vorhaben scheiterte an der Komplexität der Items. Die Anwendung der Figur-Grund-Relationen auf die Items zeigte sich bei genauerer Betrachtung als nicht angemessen. Auch die Anwendung der Methode der *Structural Information Theory* zur Erfassung der für die Annahmen relevanten Musterstruktur war anhand der verfügbaren Literatur nicht möglich. Die Definition eines Vorlagemusters als gute Gestalt ist dadurch nicht eindeutig möglich. Lediglich die Symmetrie trat als eindeutig zuordenbares Prinzip heraus. Dies gilt gleichzeitig als am stärksten mit dem Prinzip der Prägnanz verbunden, wodurch ein Einbezug der Gestaltprinzipien in das Modell zumindest insofern gelingt.

Aus der Arbeit von Royer (1977) bzw. Royer und Weitzel (1977) lässt sich aber auch ableiten, dass die Erfassung der Prägnanz des Musters als Ganzes nicht unbedingt vonnöten sein muss, um dem Wirken von Gestalt-Prozessen auf den Grund zu gehen. In den Untersuchungen bezogen die Autoren ihre Annahmen auf die Verbindung von je lediglich 2 Würfeln durch einen fehlenden Kontrast zwischen den angrenzenden Würfelkanten. In den Untersuchungen von Royer (1977), Royer und Weitzel (1977) sowie Schorr et al. (1982) zeigt sich damit eine vielversprechende Operationalisierungsmöglichkeit für die Anforderung an das Zergliedern von Gestalten. Allerdings lassen sich die Ergebnisse und Schlussfolgerungen dieser Studien nicht einfach auf den UT 10 umlegen, da dieser mehr und andersartige Würfelseiten inkludiert. Auch sind die im UT 10 enthaltenen Muster deutlich komplexer, als dass sie sich auf simple Rauten- und Streifenformen (vgl. Rozencwajg & Corroyer, 2001) reduzieren ließen. Dennoch konnte durch die Betrachtung der Herangehensweise der verschiedenen Autoren eine gewisse Betrachtungsweise übernommen werden.

Aus der Forderung nach linearer Unabhängigkeit der Vektoren in der Strukturmatrix (Fischer, 1974), ergibt sich eine besondere Problematik für die Definition der einzelnen Elementarfaktoren, denn die Anzahl an Würfeln, mit denen das Muster nachzulegen ist, bedingt viele der möglichen Elementarfaktoren. Deshalb wurde bei der Gewichtung der im Folgenden dargestellten Elementarfaktoren besondere Rücksicht darauf genommen, sofern dies möglich war.

Die ersten sechs Elementarfaktoren betreffen die Gliederung des Musters durch ein mentales Raster bzw. die Abgrenzung einzelner Würfelseiten. Sie repräsentieren das Konstrukt der Strukturierungs- und Gliederungsfähigkeit. Die kognitive Operation *rotieren* stellt eine ergänzende Fähigkeit dar. Für das Modell im Ganzen wird angenommen, dass es vollständig ist und die Schwierigkeiten der einzelnen Items genau bestimmen kann.

**H<sub>0</sub>:** Die Itemschwierigkeiten des Untertests *Analysieren und Synthetisieren* lassen sich durch die beschriebenen Elementaroperationen vollständig erklären.

Gilt das Modell mit den festgelegten Elementaroperationen, so kann darauf geschlossen werden, dass der Untertest *Analysieren und Synthetisieren* diese Fähigkeiten misst.

### **Ausgangs- und Referenzpunkt Ecke**

Aus *Testleiter-Notiz* 3 kann entnommen werden, dass das Nachlegen stets an einer Ecke beginnt. Diese Position liefert per se Informationen über die Kanten der Würfelseiten und folglich die Identität der Würfelseite. Liegen bei Eckwürfeln zusätzliche Informationen über deren Kanten auch innerhalb des Musterrahmens vor, so kann davon ausgegangen werden, dass diese einen optimalen Ausgangspunkt für das Nachlegen bilden und zudem einen Referenzpunkt für das weitere Rasterziehen darstellen können (vgl. *TLN* 6).

**H<sub>1/1</sub>:** Muster, bei denen in mindestens einer Ecke das Oberflächenelement deutlich durch Informationen über dessen Kantenverlauf abgegrenzt ist, sind einfacher zu lösen, als Muster ohne einen solchen Referenzpunkt.

Zur Erfassung dieser Kanteninformationen werden Farbwechsel im Sinne von Schorr et al. (1982, *edge cues*) verwendet. Für die Gewichtung in der *Q*-Matrix wurde eine Regel entworfen, die eine dichotome Gewichtung erlaubt. Diese ist in Tabelle 2 wiedergegeben.

**Tabelle 2:** Definition und Gewichtung von Elementarfaktor *Ecke*

<b>Ecke</b>	
Parameter:	<b>Itemmerkmal</b>
Hypothese:	<b>senkt</b> die Itemschwierigkeit
Gewichtung:	<b>dichotom</b>
Bedingung:	Mindestens drei der an den Musterrand grenzenden Kanten müssen durch das Muster oder den Musterrand sichtbar sein. Als sichtbare Kante gilt der entsprechende Farbwechsel von Weiß auf Rot.

## Kanten

Der bedeutende Unterschied in den Materialeigenschaften des UT 10 im Vergleich zu anderen *Blockdesign*-Verfahren liegt in den beiden zweifarbigen Würfelseiten, die das Muster innerhalb der Würfeloberfläche mittels horizontaler und vertikaler Linien teilen. Bei diagonal geteilten Würfelseiten ist recht deutlich zu erkennen, wo die Kanten und Eckpunkte des Würfels und damit des Rasters liegen, da die Diagonale durch diese Eckpunkte geht und die anderen Linien des Figurumrisses entlang des Rasters verlaufen. Die beiden anderen zweifarbigen Seiten hingegen, verzerrn optisch das Raster und erschweren die richtige Proportionierung des Musters, da sie den Verlauf des Figurumrisses auf einen Bereich zwischen den Rasterlinien und Kreuzpunkten versetzen. Es muss also in jedem Falle entschieden werden, ob der Figurumriss tatsächlich entlang einer Kante verläuft oder nicht.

**H<sub>1/2</sub>:** Muster, für die die zweifarbigen Würfel horizontal/vertikal geteilt und viertel Ecke benötigt werden, sind schwieriger zu lösen, als Muster ohne deren Beteiligung.

Für die *Q*-Matrix soll der Elementarfaktor entsprechend der in Tabelle 3 beschriebenen Regel einbezogen werden.

**Tabelle 3:** Definition und Gewichtung von Elementarfaktor *Kanten*

<b>Kanten</b>	
Parameter:	<b>Itemmerkmal</b>
Hypothese:	<b>erhöht</b> die Itemschwierigkeit
Gewichtung:	<b>dichotom</b>
Bedingung:	Mindestens eine Würfelseite der Sorten <i>horizontal/vertikal geteilt</i> oder <i>viertel Ecke</i> ist im Muster enthalten.

## Kreuzpunkte

Bezugspunkte für die richtige Proportionierung bzw. den Verlauf der Rasterlinien bieten nicht nur die Kantenverläufe. Kern der Raster bilden deren Kreuzpunkte. Indem sie die horizontalen Linien mit den vertikalen verbinden stabilisieren sie das Raster. Während sich bei Item 5 logischerweise (2\*1) kein Kreuzpunkt ergibt, da nur eine vertikale Linie gezogen werden muss, liegt bei allen Mustern mit 2\*2 Würfeln ein einzelner Kreuzpunkt genau in der Mitte. Bei Mustern mit 3\*3 Würfeln ergeben sich bereits 4 solcher Schnittpunkte. Angenommen wird, dass die Sichtbarkeit solcher Kreuzpunkte die Stabilität des mentalen Rasters erhöht und damit die Bearbeitung des Items erleichtert. Es hängt aber nicht nur die Anzahl der Schnittpunkte von der Anzahl der Würfel ab. Während bei 2\*2 Würfeln der eine Kreuzpunkt eben sichtbar ist oder nicht (1, 0) können bei 3\*3 Würfeln null bis vier Punkte erkennbar sein. Um einerseits die Komplexität der Matrix nicht unnötig zu erhöhen, und andererseits für eine angemessene Verteilung der über die Items hinweg zu sorgen, wird der Elementarfaktor, wie in Tabelle 4 über den Anteil an sichtbaren Kreuzpunkten erklärt. Dies soll gleichsam die Moderation des Parameters durch die Anzahl an Würfelseiten verhindern.

**H<sub>1/3</sub>:** Liegen ausreichend sichtbare Hinweise auf die Kreuzpunkte des mentalen Rasters vor, ist das Item einfacher zu lösen, wie wenn keine solche Hinweise gegeben sind.

**Tabelle 4:** Definition und Gewichtung von Elementarfaktor *Kreuzpunkte*

<b>Kreuzpunkte</b>	
Parameter:	<b>Itemmerkmal</b>
Hypothese:	<b>senkt</b> die Itemschwierigkeit
Gewichtung:	<b>dichotom</b>
Bedingung:	Mindestens 50% der möglichen Kreuzpunkte müssen sichtbar sein. Sichtbar bedeutet, dass mindestens ein erkennbarer Winkel im Figurumriss auf dem Punkt liegt. Dieser Winkel muss nicht entlang der horizontalen und vertikalen Rasterlinien verlaufen. Der Kreuzpunkt kann auch durch diagonal geteilte Würfelseiten indiziert werden.

## **dritteln**

Der bedeutende Unterschied im Raster zwischen den Mustern mit 2, 4 und 9 Würfelseiten liegt nicht nur in der Größe und Form des Rasters, sondern auch in dessen Proportionen zum Muster. Während das Rasterziehen bei 2\*1 und bei 2\*2 Würfeln lediglich ein halbieren des Bereichs erfordert, muss bei 3\*3 Würfeln der Bereich gedrittelt werden. Das bedeutet, dass nicht einfach nur mehr mentale Linien und Kreuzpunkte gezogen werden müssen, was bereits durch den Elementarfaktor *Kreuzpunkte* miterfasst wird, sondern für die Positionierung der Linien müssen komplexere Relationen zum Muster hergestellt werden. Es wird angenommen, dass das Dritteln eine größere Anforderung darstellt als das Halbieren. Indem durch *Kreuzpunkte* bereits einem bedeutenden inhaltlichen Korrelat mit der Anzahl der Würfel entsprochen wurde und dessen Gewichtung einen von der faktischen Anzahl der Würfel unabhängigen Matrizenvektor ergab, kann davon ausgegangen werden, dass der Elementarfaktor *dritteln* einen zusätzlichen Erklärungswert liefern kann. Tabelle 5 stellt die Elementaroperation schematisch dar.

**H<sub>1/4</sub>:** Das Dritteln des Musters ist schwieriger als das Halbieren des Musters.

**Tabelle 5:** Definition und Gewichtung der Elementaroperation *dritteln*

<b>dritteln</b>	
Parameter:	<b>Operation</b>
Hypothese:	<b>erhöht</b> die Itemschwierigkeit
Gewichtung:	<b>dichotom</b>
Indikator:	3*3-Matrix

## Hintergrund

Es wurden bereits spezielle Aspekte der Musterbeschaffenheit besprochen, durch die einzelne Würfelseiten identifiziert werden können. Eindeutigste Referenzpunkte bietet der Rahmen des Musters. Viele Kinder zeigen in den Instruktionsbeispielen Schwierigkeiten damit, den Rahmen adäquat für die Bestimmung des Musterbereichs zu berücksichtigen (*TLN 4*). Es wird angenommen, dass entweder die Größentransformation von der kleineren Vorlage nicht gelingt oder womöglich die gesamte weiße Fläche des Kärtchens implizit als Hintergrund des roten Musters wahrgenommen wird. Korrekte Relationen für das Muster zu setzen wäre dadurch erschwert. Dieses Problem würde durch einen hohen Anteil an weißer Fläche im Muster – insbesondere am Rand – verstärkt werden während damit gleichzeitig das Figur-Grund-Prinzip der Größe angesprochen wird. Zudem kann davon ausgegangen werden, dass eine große weiße Fläche nur wenig Hinweise auf die Relationen des Rasters liefert. Als formale Itemeigenschaft, die mehrere Aspekte der mentalen Strukturierung anspricht, wird daher der optische Hintergrund mitberücksichtigt. Um die Unabhängigkeit von der Anzahl der Würfel zu gewährleisten, wird entsprechend Tabelle 6 der Wert in Relation zur Mustergröße gesetzt.

**Tabelle 6:** Definition und Gewichtung von Elementarfaktor *Kreuzpunkte*

<b>Hintergrund</b>	
Parameter:	<b>Itemmerkmal</b>
Hypothese:	<b>senkt</b> die Itemschwierigkeit
Gewichtung:	<b>dichotom</b>
Bedingung:	Mindestens 50% der des Musters ist weiß

## Symmetrie

Die Informationen über den Beitrag von Symmetrie und Regelmäßigkeit zur Schwierigkeit der jeweiligen Muster lassen zwei entgegenlaufende Annahmen zu. Einerseits wird durch Symmetrie der roten Flächen die Gestalt gestärkt, was auf eine höhere Schwierigkeit beim Auflösen symmetrischer Muster schließen lassen würde. Andererseits jedoch sind diese Figuren leichter zu merken und insofern auch leichter nachzulegen, da nicht ständig mit der Vorlage verglichen werden muss und ein korrektes Nachlegen besser sowie schneller kontrolliert werden kann. Was aber vor allem auch für eine erleichterte Verarbeitung symmetrischer Muster spricht, ist die Annahme, dass durch die Symmetrie die im Bild erhaltene Komplexität und Information reduziert wird, in der Art wie es auch die *Structural Information Theory* umsetzt. Zudem bietet die Achse womöglich einen Orientierungspunkt für das Positionieren der einzelnen Elemente. Daher wird angenommen, dass Symmetrie des Musters erleichternd zur Lösung des Items beiträgt. Die Gewichtung in der Strukturmatrix wird in Tabelle 7 erklärt.

**H<sub>1/6</sub>: Symmetrische Muster sind einfacher zu lösen, als nicht symmetrische.**

**Tabelle 7:** Definition und Gewichtung von Elementarfaktor *Symmetrie*

Symmetrie	
Parameter:	<b>Itemmerkmal</b>
Hypothese:	<b>senkt</b> die Itemschwierigkeit
Gewichtung:	<b>dichotom</b>
Bedingung:	Mindestens eine symmetrische Achse lässt sich durch den Mittelpunkt des Musters ziehen – entweder diagonal (45° oder 135°), horizontal mittig oder vertikal mittig.

## Rotation

Anders als die einfarbigen Würfel, können die zweifarbigen Würfel durch eine andere Position auch eine andere Identität erhalten. Um dennoch auf die korrekte der alternativen Würfelseiten schließen zu können, müssen zweifarbiges Muster mental rotiert werden. Auch um umgekehrt die korrekte Positionierung eines ausgewählten Würfels bestimmen zu können, muss mental ein bereits rotiertes Abbild davon vorliegen, um dies mit dem Ergebnis zu

vergleichen – damit ist nicht die manuelle Rotation des Würfels an seinem designierten Platz gemeint. Tabelle 8 zeigt die Umsetzung für die *Q*-Matrix.

**Tabelle 8:** Definition und Gewichtung der Elementaroperation *rotieren*

rotieren	
Parameter:	<b>Operation</b>
Gewichtung:	<b>Anzahl</b>
Indikator:	individuell ausgerichtete zweifarbiges Würfelseite

## 14. Das Modell in Form der *Q*-Matrix

Nachdem alle Items anhand der in den Tabellen 2 bis 7 beschriebenen Regeln gewichtet wurden, ergibt sich die Strukturmatrix für das LLTM, welche in Tabelle 9 abgebildet ist.

**Tabelle 9:** vollständige Q-Matrix mit den sieben Elementarfaktoren *Ecke*, *Kanten*, *Kreuzpunkte*, *dritteln*, , *Hintergrund*, *Symmetrie* und *rotieren*

<b><i>Q</i>-Matrix</b>							
<b>Item</b>	<b>Ecke</b>	<b>Kanten</b>	<b>Kreuzpunkte</b>	<b>dritteln</b>	<b>Hintergrund</b>	<b>Symmetrie</b>	<b>rotieren</b>
<b>1</b>	1	1	0	0	0	1	1
<b>2</b>	1	0	1	0	1	1	1
<b>3</b>	0	0	0	0	0	1	2
<b>4</b>	0	0	1	1	0	0	4
<b>5</b>	1	0	1	0	0	0	0
<b>6</b>	1	1	0	0	1	0	2
<b>7</b>	1	1	1	0	0	0	2
<b>8</b>	0	0	1	0	0	0	2
<b>9</b>	0	0	0	1	1	0	2
<b>10</b>	1	1	1	1	1	1	4
<b>11</b>	0	0	1	0	0	1	4
<b>12</b>	0	1	1	1	1	0	4
<b>13</b>	0	1	0	1	1	1	3
<b>14</b>	1	1	0	1	1	0	4
<b>15</b>	1	1	1	1	1	0	4
<b>16</b>	0	0	1	1	0	1	4
<b>17</b>	1	1	0	1	1	1	4
<b>18</b>	0	1	1	1	0	0	3

# VI. Modellprüfung

---

Für die folgenden Berechnungen wurde ein Auszug aus dem Datensatz für die Normierung des AID 3 herangezogen. Dieser wurde der Verfasserin von AID 3 Autor Klaus D. Kubinger zur Verfügung gestellt. Eine ausführliche Beschreibung des Datensatzes findet sich bei Kubinger & Holocher-Ertl (in Druck).

## 15. Stichprobe und Datensatz

Der Datensatz umfasst insgesamt 1130 Testpersonen mit einem durchschnittlichen Alter von 10,8 Jahren. Die Daten wurden in sowohl in Österreich (864) als auch in Deutschland (266) erhoben. Insgesamt 238 Testpersonen, also 21% der Stichprobe, haben eine nichtdeutsche Muttersprache. Die Geschlechterverteilung ist mit 539 männlichen und 591 weiblichen Personen ausgewogen.

## 16. Qualität der *Q*-Matrix

Bei Betrachtung von Tabelle 9 zeigt sich eine angemessene Differenziertheit der *Q*-Matrix. Wenngleich aufgrund des Verhältnisses zwischen der Anzahl der Elementarfaktoren und der Anzahl der Items nicht alle möglichen Elementarfaktorkombinationen in der Matrix vertreten sein können, so ergibt sich doch für jedes Item eine individuelle Elementarfaktorkombination. Dies spricht für die Qualität der Matrix, da dies eine gut differenzierte Itemparameterschätzung ermöglicht. Es sollte sich für jedes Item ein individueller Schätzwert ergeben. Bei genauer Betrachtung der Vektoren zeigen sich diese als weitgehend linear unabhängig. Lediglich für *rotieren* und *dritteln* wurde bereits [s6]theoretisch ein möglicher Zusammenhang aufgezeigt, da der maximal mögliche Wert für *rotieren* von der Größe der Vorlage abhängt, welche durch *dritteln* mit repräsentiert wird. Tatsächlich entspricht dies der in Tabelle 10 abgebildeten Korrelationsmatrix. Der mittlere Zusammenhang, der hier für *rotieren* und *dritteln* angegeben wird, ist damit erklärbar. Abgesehen davon sind die beiden Operationen jedoch inhaltlich unabhängig voneinander. Deshalb und weil keiner der beiden Elementarfaktoren mit einem anderen bedeutsam korreliert, wird von deren Ausschluss abgesehen. Die Matrix hat vollen Rang.

**Tabelle 10:** Korrelationsmatrix mit den sieben Elementarfaktoren *Ecke*, *Kanten*, *Kreuzpunkte*, *dritteln*, , *Hintergrund*, *Symmetrie* und *rotieren*

lineare Zusammenhänge der <i>Q</i> -Matrix Vektoren						
	Ecke	Kanten	Kreuzpunkte	dritteln	Hintergrund	Symmetrie
R.Kanten	0,45					
R.Kreuzpunkte	-0,11	-0,25				
dritteln	-0,22	0,33	-0,03			
Hintergrund	0,33	0,45	-0,34	0,45		
Symmetrie	0,00	-0,10	-0,20	-0,10	0,00	
rotieren	-0,26	0,28	0,13	0,72	0,26	0,07

## 17. Modellprüfungen

Wie bereits in Abschnitt IV dargestellt, ist die Geltung des *Rasch*-Modells für die Anwendung des *LLTM* Voraussetzung. Die sich aus der Berechnung ergebenden Itemparameter stellen die Grundlage für den Vergleich der beiden Modelle dar. Wenngleich entsprechend Kubinger und Holocher-Ertl (in Druck) von der Eindimensionalität des Itempools ausgegangen werden kann, wird im Folgenden zur Absicherung und genaueren Betrachtung die Konformität der Daten mit dem *Rasch*-Modell wiederholt analysiert. Dabei wird auf problematische Items geachtet und die Möglichkeit, einzelne Items zugunsten der Modellpassung auszuscheiden, in Betracht gezogen. Im Anschluss wird der modellkonforme Itempool für die Berechnung des *LLTM* herangezogen. Für die statistische Analyse wird das Softwareprogramm R (R Core Team, 2013) mit dem Paket eRm (Mair, Hatzinger & Maier, 2013) verwendet.

### 17.1 Prüfung der *Rasch*-Modell-Konformität aller Items

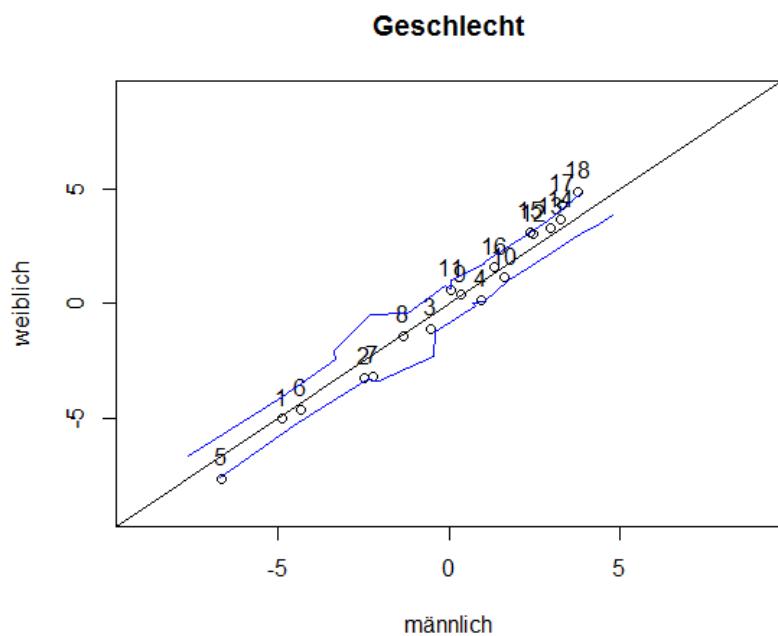
Ausgehend von den 18 zweidimensionalen Items des UT 10 wurde das *RM* berechnet und mittels *LRT* überprüft. Als Teilungskriterien wurden hierfür, entsprechend Kubinger und Holocher-Ertl (in Druck) das Geschlecht der Testpersonen sowie die Staatszugehörigkeit herangezogen. Die Sprache konnte aufgrund einer unangemessenen Verteilung nicht als Teilungskriterium herangezogen werden. Das Trennkriterium Score wurde bewusst nicht gewählt, da dies aufgrund des sogenannten *Waldenmair-Phänomens* zu einem methodischen Artefakt führen würde. Im Manual des AID 2 findet sich eine ausführliche Argumentation hierzu (Kubinger und Wurst, 2000; genauer bei Kubinger, 1989, und Draxler, 2002). Für die Berechnungen anhand mehrerer Trennkriterien wurde das Risiko 1. Art, ebenfalls entsprechend Kubinger und Holocher-Ertl (in Druck), mit 0,01 festgelegt. Die Ergebnisse für

den *LRT* je Trennkriterium werden in Tabelle 11 dargestellt. Da in beiden Fällen der empirische  $\chi^2$ -Wert kleiner als der kritische Wert ist, kann von Modellgültigkeit ausgegangen werden.

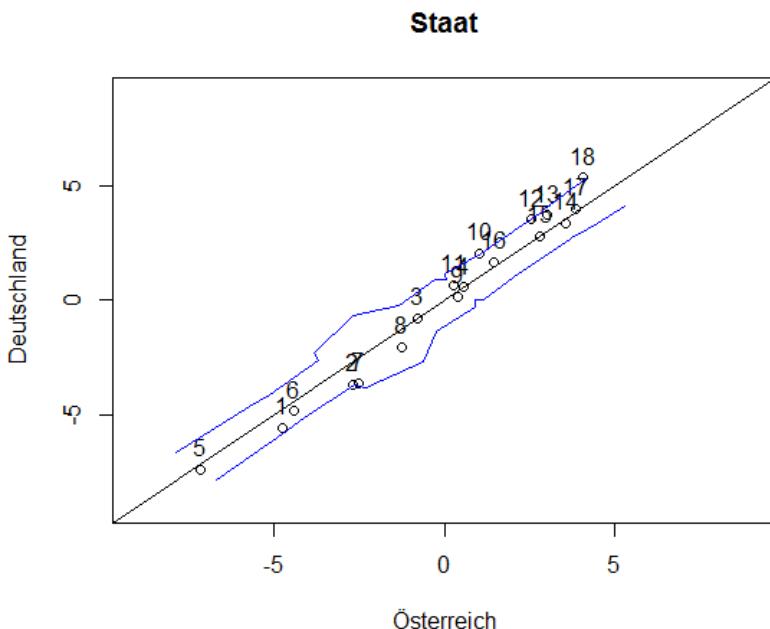
**Tabelle 11:** Ergebnisse der RM-Prüfung anhand des LRT nach Andersen für die Trennkriterien Geschlecht und Staat

<b>Likelihood Ratio Test für das Rasch-Modell</b>				
Trennkriterium	$\chi^2$	$\chi^2_{\text{krit}} (\alpha = 0,01)$	$df = k - 1$	p
Geschlecht	21,398	33,409	17	0,209
Staat	31,845	33,409	17	0,016

Einer differenzierten Betrachtung des Ergebnisses sollen Abbildung 5 und Abbildung 6 dienen. Aus ihnen geht die Passung der jeweiligen Itemparameterschätzungen je Gruppe für die Trennkriterien hervor. Insgesamt zeigt sich für den UT 10 eine gute Modellkonformität. Item 18 zeigt sich jedoch bei beiden Trennkriterien, stärker aber bei der Staatszugehörigkeit, als leicht abweichend. Bei Betrachtung des Items wird kein inhaltlicher Grund dafür deutlich. Die Abweichung ist jedoch gering und der *LRT* spricht für eine ausreichende Modellkonformität. Die Berechnung des *LLTM* wird jedenfalls als zulässig erachtet.



**Abbildung 5:** graphische Kontrolle des RM mit allen Items unter Einbezug des Konfidenzbandes ( $\alpha = 0,01$ ) für das Trennkriterium Geschlecht



**Abbildung 6:** graphische Kontrolle des *RM* mit allen Items unter Einbezug des Konfidenzbandes ( $\alpha = 0,01$ ) für das Trennkriterium Staat

## 17.2 Prüfung der Geltung des *LLTM* mit allen Items

Wie aus Tabelle 12 ersichtlich, zeigt der *Likelihood Ratio Test* eine hohe, signifikante Abweichung der Itemparameter im *LLTM* von den Itemparametern im *Rasch-Modell* auf. Deren Korrelationskoeffizient ist zwar mit 0,91 relativ hoch, der empirische  $\chi^2$ -Wert des *LRT* ist jedoch um Vieles größer, als der kritische  $\chi^2$ -Wert. Die Nullhypothese der Nebenbedingungen muss daher verworfen werden.

**Tabelle 12:** Ergebnis der *LLTM*-Prüfung mittels *LRT* nach Andersen

<b><i>Likelihood Ratio Test für das LLTM</i></b>				
<b>Parameter</b>	$\chi^2$	$\chi^2_{\text{krit}} (\alpha = 0,05)$	$df = k - m - 1$	<i>p</i>
7	595,971	18,307	10	0,000

Die graphische Modellkontrolle (Abbildung 7) zeigt deutlich, dass vor allem die Schwierigkeiten der Items 6 und 18 nicht angemessen durch das *LLTM* vorhergesagt werden können, während für die restlichen Itemparameter relativ übereinstimmende Schätzungen erzielt wurden. Diese beiden Punkte liegen am weitesten von der Linie entfernt.

### graphische Modellkontrolle des LLTM

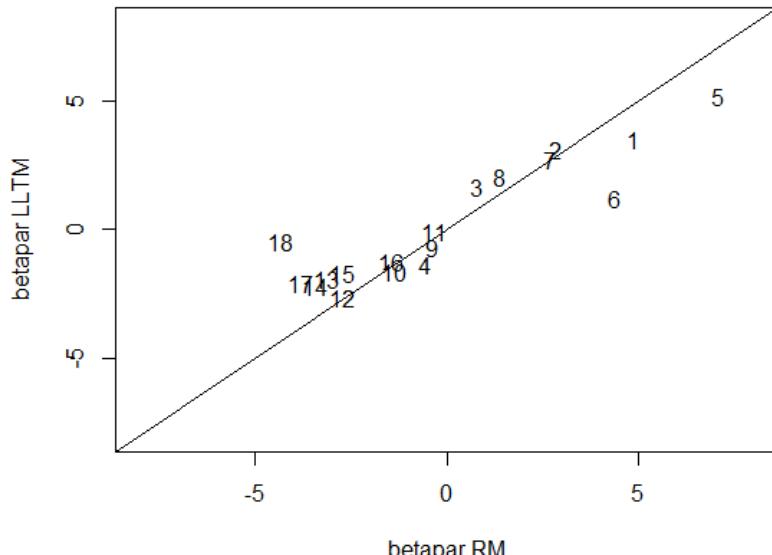


Abbildung 7: graphische Kontrolle des LLTM mit allen 18 Items

Die Parameterschätzungen für die jeweiligen Elementarfaktoren sind in Tabelle 13 wiedergegeben. Die Signifikanz der Parameter wurde mittels Konfidenzintervall ( $KI$ ) bei  $\alpha = 0,05$  überprüft. Bis auf *Symmetrie* scheinen alle Elementarfaktoren signifikant zur Vorhersage der Schwierigkeit beizutragen. Die Vorzeichen der jeweiligen Parameter stimmen mit den aufgestellten Hypothesen überein. Positive Werte sprechen für ein Erleichtern der Aufgabe während negative Werte dafür stehen, dass die kognitive Operation oder die Itemeigenschaft, das Bearbeiten der Aufgabe erschweren. Dennoch deuten sowohl das Ergebnis des LRT, als auch die schlechten Schätzungen für die Items 6 und 18, auf eine mangelnde Beschreibung der für den UT 10 erforderlichen Denkprozesse durch die hier beschriebenen Elementarfaktoren.

Tabelle 13: Schätzungen der Basisparameter mit Schätzfehler ( $SE$ ) im LLTM mit allen Items

Elementarfaktor	Parameter	SE	$KI_{\alpha=0,05}$
Ecke	0,951	0,084	sign.
Kanten	-0,239	0,091	sign.
Kreuzpunkte	0,481	0,078	sign.
dritteln	-1,168	0,145	sign.
Hintergrund	-1,080	0,092	sign.
Symmetrie	0,105	0,071	n. sign.
rotieren	-1,113	0,057	sign.

Für die Auffälligkeiten der Items 6 und 18 ist inhaltlich kein Grund ersichtlich. Da aber Item 18 bereits in der graphischen Kontrolle des *Rasch*-Modells auffällig wurde und sich hier nun ebenfalls als abweichend präsentiert, kann vermutet werden, dass möglicher Weise trotz insignifikanter *LRTs* eine Verletzung der *Rasch*-Modell-Konformität durch dieses Item bestehen. Daher soll das *LLTM* unter Ausschluss von Aufgabe 18 neu berechnet werden. Hierdurch wird auch eine erneute Berechnung des *Rasch*-Modells erforderlich.

### 17.3 Prüfung der *Rasch*-Modell-Konformität nach Ausschluss von Item 18

Auch nach Ausschluss von Item 18 hat das *Rasch*-Modell Gültigkeit (Tabelle 14). Es zeigt sich sogar, dass insbesondere beim Trennkriterium Staat nun von einem besseren Ergebnis im Sinne der Passung gesprochen werden kann. Eine Betrachtung auf Itemebene soll daher ausgelassen und stattdessen gleich zur neuen Prüfung des *LLTM* übergegangen werden.

**Tabelle 14:** Ergebnisse der *RM*-Prüfung nach Ausschluss von Item 18 anhand des *LRT* nach Andersen für die Trennkriterien Geschlecht und Staat

<i>Likelihood Ratio Test</i> für das <i>Rasch</i> -Modell ohne Item 18				
Trennkriterium	$\chi^2$	$\chi^2_{\text{krit}} (\alpha = 0,01)$	$df = k - 1$	$p$
Geschlecht	19,310	31,999	16	0,253
Staat	25,555	31,999	16	0,061

### 17.4 Prüfung der Geltung des *LLTM* nach Ausschluss von Item 18

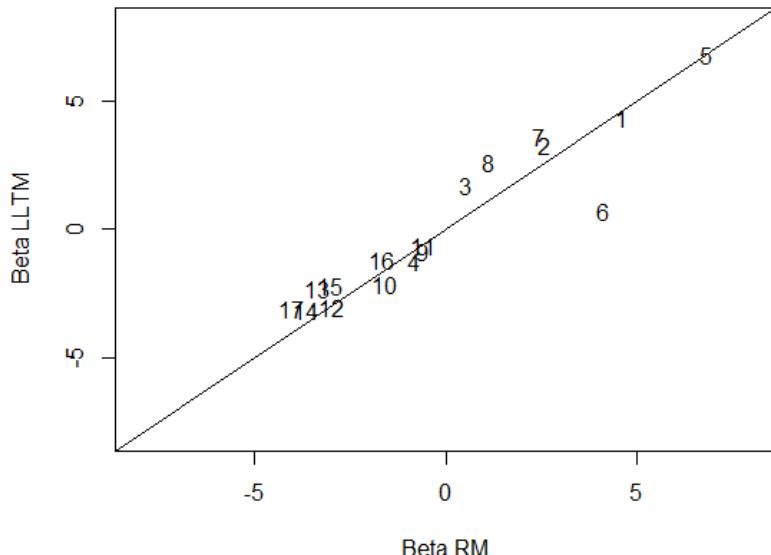
Der Ausschluss von Item 18 führte zu einer deutlich besseren Passung zwischen *LLTM* und *Rasch*-Modell (Tabelle 15). Allerdings ist der empirische  $\chi^2$ -Wert nach wie vor um Vieles größer, als es der kritische  $\chi^2$ -Wert für ein nicht signifikantes Ergebnis erlaubt. Das *LLTM* erreicht also auch nach Ausschluss von Item 18 keine Gültigkeit [S7]. Auch wenn die Zulässigkeit des Vergleichs zweier signifikanter Testergebnisse hinterfragt werden muss, bietet die verbesserte *Rasch*-Modell-Konformität des reduzierten Itempools eine plausible Erklärung für die geringere Abweichung des *LLTM*.

**Tabelle 15:** Ergebnisse der LLTM-Prüfungen mittels LRT nach Andersen vor und nach Ausschluss von Item 18

<b>Likelihood Ratio Test für das LLTM</b>					
Items	Parameter	$\chi^2$	$\chi^2_{\text{krit}} (\alpha = 0,05)$	$df = k - m - 1$	p
1-18	7	595,971	18,307	10	0,00
1-17	7	251,924	16,919	9	0,00

Abbildung 8 zeigt deutlich, dass der Schwierigkeitsparameter von Item 6 nach wie vor nicht adäquat durch die festgelegten Basisparameter vorhergesagt werden kann. Der Punkt scheint sogar noch weiter von der Linie abzuweichen, als zuvor. Die Schätzungen für die anderen Itemparameter des Rasch-Modells sind hingegen relativ akkurat, wobei im mittleren Schwierigkeitsbereich die schletesten Übereinstimmungen liegen. Der Korrelationskoeffizient für die Itemparameter der beiden Modelle für 17 der Aufgaben ist mit 0,94 nur geringfügig höher als der für die Modelle mit allen Items.

#### graphische Modellkontrolle des LLTM (ohne 18)



**Abbildung 8:** graphische Kontrolle des LLTM nach Ausschluss von Item 18

Kritisch zu betrachten ist das Ergebnis der Basisparameterschätzungen (Tabelle 16), wonach *Kanten*, anders als in der vorherigen Schätzung, einen Leichtigkeitsfaktor darstellt. Allerdings leistet der Elementarfaktor diesmal auch keinen signifikanten Beitrag für die Itemparameterschätzung. Die Parameterschätzungen für die anderen Elementarfaktoren haben

sich zwar, wie zu erwarten, auch stark verändert, doch stehen diese nach wie vor in Einklang mit den entsprechenden Hypothesen ( $H_{1/1}, H_{1/3}, H_{1/4}, H_{1/5}, H_{1/7}$ ).

**Tabelle 16:** Schätzungen der Basisparameter Schätzfehler ( $SE$ ) im  $LLTM$  nach Ausschluss von Item 18

Elementarfaktor	Parameter	SE	$KI_{\alpha=0,05}$
Ecke	0,852	0,091	sign.
Kanten	0,135	0,099	n. sign.
Kreuzpunkte	0,985	0,090	sign.
dritteln	-0,599	0,155	sign.
Hintergrund	-1,911	0,109	sign
Symmetrie	0,057	0,076	n. sign.
rotieren	-1,658	0,073	sign.

Für eine abschließende Betrachtung werden die Itemparameter des  $LLTM$  jenen des Rasch-Modells gegenübergestellt (Tabelle 17).

**Tabelle 17:** Gegenüberstellung der Itemparameter ( $\sigma_i$ ) des  $RM$  und des  $LLTM$  nach Ausschluss von Item 18 mit dazugehörigen Schätzfehlern ( $SE$ )

Itemparameter in $RM$ und $LLTM$ nach Ausschluss von Item 18				
Item	$\sigma_i^{(RM-18)}$	SE	$\sigma_i^{(LLTM-18)}$	SE
1	<b>4,641</b>	0,225	<b>4,171</b>	0,138
2	<b>2,590</b>	0,198	<b>3,110</b>	0,182
3	<b>0,552</b>	0,202	<b>1,525</b>	0,166
4	<b>-0,807</b>	0,225	<b>-1,463</b>	0,295
5	<b>6,850</b>	0,323	<b>6,622</b>	0,124
6	<b>4,142</b>	0,467	<b>0,545</b>	0,202
7	<b>2,452</b>	0,196	<b>3,441</b>	0,151
8	<b>1,136</b>	0,220	<b>2,453</b>	0,130
9	<b>-0,602</b>	0,162	<b>-1,042</b>	0,227
10	<b>-1,615</b>	0,242	<b>-2,330</b>	0,348
11	<b>-0,572</b>	0,188	<b>-0,807</b>	0,283
12	<b>-2,980</b>	0,197	<b>-3,238</b>	0,330
13	<b>-3,367</b>	0,200	<b>-2,508</b>	0,306
14	<b>-3,675</b>	0,204	<b>-3,371</b>	0,328
15	<b>-3,010</b>	0,202	<b>-2,386</b>	0,316
16	<b>-1,687</b>	0,203	<b>-1,406</b>	0,325
17	<b>-4,046</b>	0,232	<b>-3,315</b>	0,352

## VII. Diskussion

---

Das Ziel der vorliegenden Arbeit, nämlich die kognitiven Operationen für die Lösung von Aufgaben des Tests *Analysieren und Synthetisieren*, vollständig zu identifizieren und bestätigen, konnte nicht erreicht werden. Das erstellte Modell kann den Lösungsprozess im UT 10 nicht hinreichend erklären. Wenngleich große Übereinstimmungen in den vielzähligen Beschreibungen der mentalen Prozesse und ausschlaggebenden Mustereigenschaften bei *Blockdesign*-Aufgaben und Tests zur Feldunabhängigkeit gefunden wurden, konnten diese hier nicht passend in eine Strukturmatrix gefasst werden. Deren Resultat enthält neben zwei abgrenzbaren kognitiven Operationen auch 4 formale Aspekte der Itemmuster, welche alle als Einflussfaktoren des Strukturierens und Gliederns hypothetisiert wurden.

Im ersten Durchgang der Berechnungen war ein Item (18) im Itempool enthalten, das zu weit außerhalb des durch das *Rasch*-Modell definierten Itemuniversums liegt, für welches das *LLTM* den Anspruch auf Gültigkeit erhebt. Wie besprochen ist die *Rasch*-Modell-Konformität notwendige Voraussetzung für die Geltung des *LLTM*. Das Modell wurde daher mit der Einschätzung eines quasi „fremden“ Items überfordert. Das Entfernen von Item 18 aus dem zu beschreibenden Itempool führte, wie zu erwarten, zu einer geringeren Abweichung, welche jedoch ebenfalls zu groß für die Modellgültigkeit ist. In Anlehnung an Kubinger (1979) kann argumentiert werden, dass die mangelnde Modellpassung womöglich durch die hohe Sensitivität des *Likelihood Ratio Tests* bedingt ist. Betrachtet man die auffällig starke Abweichung der Schwierigkeitsschätzung für Item 6, so wird allerdings deutlich, dass die hypothetisierten Basisparameter des *LLTM* nicht wie gefordert alle Items im Itemuniversum gleichermaßen genau beschreiben können. Würden aus diesem Modell Konstruktionsregeln abgeleitet, so könnte nicht davon ausgegangen werden, dass die tatsächliche Schwierigkeit der damit erstellten neuen Items den Erwartungen entsprechen wird.

Dass eine einzelne Kombination der Elementarfaktoren zu einer derart größeren Abweichung führt, spricht dafür, dass entweder eine unentdeckte Beziehung zwischen den Elementarfaktoren besteht, oder mindestens ein relevanter Elementarfaktor existiert, der nicht in das Modell integriert wurde.

Bei der Integration zusätzlicher Basisparameter ist jedoch stets auf die Sparsamkeit des Modells zu achten. Besonders bei einem recht kleinen Itempool, wie dem um die dreidimensionalen Items reduzierten UT 10, sollte jeder Basisparameter bereits vor der Aufnahme in das Modell auf seine Sinnhaftigkeit, auch in Bezug auf die anderen Elementarfaktoren, hin geprüft werden.

Aufgrund der Größe des Itempools wurde auch das Dichotomisieren der Elementarfaktoren gar nicht erst in Betracht gezogen, sondern jeweils nach Möglichkeit gleich dichotom gewichtet. Dadurch gingen viele Informationen verloren oder wurden gar nicht erst berücksichtigt. Die Erweiterung durch die individuelle Berücksichtigung jeder Faktorstufe könnte jedoch eine Möglichkeit darstellen, die Hypothesen durch ein anderes Gewichtungsschema erneut zu prüfen. Die Anzahl der Vektoren würde damit allerdings in jedem Falle stark ansteigen.

Dass eine Erweiterung des Modells nötig scheint, spricht für eine hohe Komplexität des Denkprozesses, der für die erfolgreiche Bearbeitung der Aufgaben im UT 10 nötig ist.

Können bestimmte Annahmen über Interaktionen der Elementarfaktoren untereinander aufgestellt werden, so könnte diesen durch weitere Elementarfaktordefinitionen entsprochen werden. Es wäre beispielsweise möglich, statt einzelnen Elementarfaktoren bestimmte Kombinationen, also das gemeinsame Auftreten bestimmter Ausprägungen, zu werten. Ein erster Hinweis auf eine solche wechselseitige Beeinflussung liegt bereits in Form der weiter oben abgebildeten Korrelationsmatrix (Tabelle 10) vor. Wie berichtet, zeigte sich ein mittlerer linearer Zusammenhang zwischen den Anforderungen für die Operationen *dritteln* und *rotieren*. Beide scheinen durch die Anzahl der Würfel je Item mitbedingt zu sein, während für die anderen Vektoren eine davon weitgehend unabhängige Gewichtung gelungen ist. Der Preis für diese Unabhängigkeit ist Information, die auf dem Weg der Umrechnung der gemessenen Werte, erst in Prozente und dann wieder in ein rangskaliertes Maß, verloren ging. Tatsächlich stellte der Umstand dieser offensichtlichen Interaktionen die größte Herausforderung an die Erstellung der Matrix dar. Eine unangemessene Gewichtung der einzelnen Elementarfaktoren scheint als Grund für die große Modellabweichung insofern aber vergleichsweise unwahrscheinlich, als dass sie für 16 der 17 Items als weitgehend angebracht wirkt.

Die zumindest annähernd gute Einschätzung für diese 16 Items kann nicht über die mangelnde Modellgültigkeit hinwegtäuschen. Betrachtet man die einzelnen

Elementarfaktoren für sich, so zeigt sich aber doch noch ein wenigstens halbwegs zufriedenstellendes Ergebnis. Auch wenn das erstellte Modell den Lösungsprozess nicht genau abbildet, so konnte doch für vier der sechs Strukturierungs- und Gliederungsparameter sowie für die mentale Rotation dokumentiert werden, dass sie zu einer groben Vorhersage der Itemschwierigkeiten beitragen können. Dies spricht für die Konstruktvalidität des Tests. Entsprechend der aufgestellten Hypothesen tragen a) das Vorhandensein mindestens eines leicht identifizierbaren Würfels an einer Eckposition und b) die Erkennbarkeit der Rasterkreuzpunkte zur Leichtigkeit der Items bei, während c) die Erforderlichkeit der Rasterproportionierung mittels Drittteilung und d) der Anteil an weißer Fläche sowie e) der nötige Rotationsaufwand die Items schwieriger machen. Berücksichtigt man die Ergebnisse von Baker (1993), so ist die Zulässigkeit dieser Interpretation zu hinterfragen.

Es wird aus diesem Ergebnis dennoch gefolgert, dass das hier erstellte Modell wenigstens als eine gute Ausgangsbasis für weitere Untersuchungen dienen kann. Bei einer neuerlichen Anwendung des *LLTM* wäre unter anderem ein größerer Itempool wünschenswert. Liegen der *Q*-Matrix tatsächlich bedeutsame Interaktionen zugrunde, so könnte durch die Steigerung der maximal möglichen individuellen Kombinationen in der Matrix, deren Aufdecken erleichtert werden. Zudem ergäbe sich durch eine erhöhte Anzahl an Items auch mehr Raum für die potentielle Inklusion zusätzlicher Parameter.

Zwar kann – gerade in Anbetracht der aktuellen Neuauflage des AID – nicht davon ausgegangen werden, dass in naher Zukunft weitere zweidimensionale Items für den UT 10 konstruiert und in das Verzweigungsschema des *banchet testings* integriert werden. Es sollte jedoch in diesem Zusammenhang die Möglichkeit in Betracht gezogen werden, die hier erarbeiteten Kriterien auf deren Anwendbarkeit für die dreidimensionalen Items des UT 10 hin zu überprüfen, oder diese entsprechend anzupassen. Dabei wäre allerdings mit einer gesteigerten Komplexität der Matrix zu rechnen. Eine Erweiterung des Anwendungsbereiches der vorliegenden Definitionen je Elementarfaktor sollte, aufgrund deren allgemein gehaltener Formulierung, durch die Anpassung der Gewichtungsregeln möglich sein.

Der empirische Teil dieser Arbeit lieferte nicht die gewünschte Bestätigung, dass die identifizierten Elementarfaktoren den Lösungsprozess vollständig abbilden. Rückschlüsse auf Teilaspekte des Lösungsprozesses wurden aus den Ergebnissen dennoch gezogen. Abschließend bleibt damit zu sagen, dass die vorliegende Arbeit nicht nur durch den empirischen Teil, sondern vor allem auch durch das Zusammenführen relevanter Literatur zu bedeutenden

Schlussfolgerungen über das im Untertest *Analysieren und Synthetisieren* erfasste Konstrukt, ein wertvoller Beitrag zu dessen Validierung leisten konnte. Es bleibt die Annahme, dass zumindest fünf der hier hypothetisierten Basisparameter zur groben Vorhersage der Itemschwierigkeiten beitragen können. Zulässige Aussagen über deren individuellen Beitrag sind jedoch erst durch deren Einbettung in ein passendes Modell möglich, was hier nicht gelang.

## VIII. Zusammenfassung

---

Um jene mentalen Operationen zu identifizieren, welche für die Lösung der Aufgaben im Untertest *Analysieren und Synthetisieren* des AID 3 erforderlich sind, wurde sich thematisch mit dem Konstrukt der Strukturierungs- und Gliederungsfähigkeit beschäftigt. Den theoretischen Rahmen hierfür bildeten die Ausführungen zu Thurstone's (1944) Faktor *Freiheit von Gestaltbindung* und zur Feldabhängigkeit/Feldunabhängigkeits-Dimension. Es wurde die Schlussfolgerung gezogen, dass es sich bei beidem um identische Konstrukte handelt, die sich weitgehend mit der operanten Definition des UT 10 decken. Aus den Theorien der Gestaltpsychologie sowie aus der Beschreibung experimenteller Studien mit verschiedenen Arten von *Blockdesign*-Aufgaben wurden einerseits Annahmen über die Verarbeitung bestimmter Musterinformationen abgeleitet und andererseits methodische Herangehensweisen für die Modellerstellung berücksichtigt, wenngleich erstere nicht in gewünschtem Maße zum Einsatz kamen.

In mehreren Untersuchungen zeigte sich, dass Testpersonen bei der Bearbeitung von *Blockdesign*-Aufgaben verschiedene Strategien anwenden (Beuscart-Zéphir & Beuscart, 1988; Kurian et al., 1987; Rozencwajg, 1991; Rozencwajg & Corroyer, 2001; Schorr et al., 1982). Ein analytisch-strukturierendes Vorgehen wurde hierbei mit einer kurzen Bearbeitungszeit und dem Bilden eines mentalen Rasters in Verbindung gebracht. Letzteres scheint bei einer globalen Bearbeitungsweise nicht im selben Maße erforderlich, da durch Versuch und Irrtum eine schrittweise Anpassung des Musters an die Vorlage vorgenommen wird. Dies geht mit einer höheren Bearbeitungszeit einher. Es wurde argumentiert, dass durch die Zeitbegrenzung für die Bearbeitung der Items im Untertest *Analysieren und Synthetisieren*, allein das analytisch-strukturierende Vorgehen zu einer Lösung der Items führen kann.

Da mittels *Linear Logistischem Testmodell (LLTM)* empirisch geprüft werden kann, ob von einem Test tatsächlich die hypothetisierten kognitiven Operationen – in diesem Fall die an der Rasterbildung beteiligten elementaren Prozesse – erfasst werden, war dessen Anwendung das designierte Ziel der Arbeit. Um den Anforderungscharakter der Items an die Strukturierungs- und Gliederungsfähigkeit zu erfassen, wurden die Elementarfaktoren *Ecke*, *Kanten*, *Kreuzpunkte*, *dritteln*, *Hintergrund* und *Symmetrie* in eine geeignete Strukturmatrix für die

Prüfung des *Linear Logistischen Testmodells* gefasst. Beobachtungen aus der praktischen Erfahrung mit dem Untertest wurden hierfür ergänzend berücksichtigt. Den Elementarfaktoren liegt die Annahme zugrunde, dass die mentale Gliederung des Musters erschwert wird, wenn die Bildinformationen dieser Struktur zuwiderlaufen, und dass hingegen rasterkongruente Bildinformationen das mentale Gliedern der Muster erleichtern. Zudem wurde nach den Ergebnissen von Schorr et al. (1982) sowie Rozencwajg & Corroyer (2001) die mentale Rotation als zusätzlicher Schwierigkeitsfaktor definiert und in die Matrix aufgenommen. Als größte Schwierigkeit der Umsetzung stellte sich die Moderation der einzelnen Elementarfaktoren durch die Anzahl der Würfel, die für das Nachlegen der Muster benötigt werden, heraus. Durch die Anpassung der Gewichtungen konnte eine weitgehend linear unabhängige Strukturmatrix erarbeitet werden.

Als Berechnungsgrundlage diente ein Datensatz mit  $n = 1130$  Testpersonen, der im Rahmen von Normierungserhebungen für den AID 3 gewonnen wurde. Die Prüfung des *Rasch*-Modells als Voraussetzung für die Anwendung des *LLTM* ergab einen homogenen Itempool, wobei Item 18 leicht auffällig war. Eine erste Prüfung des *LLTM* anhand des Andersen *Likelihood Ratio Tests* zeigte eine große Abweichung der durch das *LLTM* geschätzten Itemparameter. Da neben der besonders ungenauen Schätzung für das Item 6 auch Item 18 abermals als auffällig hervorging, wurden die Berechnungen unter Ausschluss dieser Aufgabe erneut durchgeführt. Der zweite Durchgang ergab eine deutlich geringere Abweichung der Itemparameterschätzungen der beiden Modelle. Die Signifikanzprüfung fordert dennoch die Verwerfung Annahme, dass durch das *LLTM* die Itemschwierigkeiten im UT 10 hinreichend genau vorhergesagt werden können. Die erstellte Strukturmatrix kann entgegen der Nullhypothese das Konstruktional der Items nicht adäquat beschreiben. Die Betrachtung der graphischen Modellkontrolle zeigt, dass für 16 der 17 Items die Itemparameterschätzungen relativ weit übereinstimmen, während Item 6 nach wie vor weit ab der Linie liegt. Es wird im *LLTM* viel zu schwierig eingeschätzt.

Als mögliche Gründe für die mangelnde Übereinstimmung der Schätzungen wurden das Fehlen relevanter Operation in der Strukturmatrix sowie eine Interaktion zwischen den einzelnen Elementarfaktoren diskutiert. Das, in Anbetracht der Ergebnisse, nicht ausreichend von der Anzahl an Würfeln unabhängige Gewichtungsschema wurde als mögliche Fehlerquelle angeführt. Trotz der kritischen Betrachtung des entworfenen Modells im Ganzen wurden einzelne relevante Aspekte der Strukturierungsfähigkeit sowie Rotationsleistung mit

der Bearbeitung des UT 10 in Verbindung gebracht. Für die Faktoren *Ecke*, *Kreuzpunkte*, *dritteln*, *Hintergrund* und *rotieren* ergab die Berechnung signifikante Basisparameter in Richtung der aufgestellten Hypothesen. Dies spricht dafür, dass a) das Vorhandensein mindestens eines leicht identifizierbaren Würfels an einer Eckposition und b) die Erkennbarkeit der Rasterkreuzpunkte zur Leichtigkeit der Items beitragen, während c) der Anteil an weißer Fläche, d) die Erforderlichkeit der Rasterproportionierung mittels Drittteilung und e) der nötige Rotationsaufwand die Items schwieriger machen. Eine solche Interpretation ist allerdings mit Vorsicht zu begegnen.

Ausblickend wurde der mögliche Nutzen des erarbeiteten Modells für die weitere Auseinandersetzung mit der Thematik besprochen. Es wird angenommen, dass eine Erweiterung der Matrix zu passenderen Itemparameterschätzungen führen würde und auch unter Einbezug der dreidimensionalen Items des Untertests möglich wäre.

## IX. Literaturverzeichnis

---

- Andersen, E. B. (1973). A Goodness of Fit Test for the Rasch Model. *Psychometrika*, 38(1), 123–140.
- Asch, S. E. & Witkin, H. A. (1948a). Studies in Space Orientation: I. Perception of the Upright with Displaced Visual Fields. *Journal of experimental psychology*, 38(3), 325–37.
- Asch, S. E. & Witkin, H. A. (1948b). Studies in Space Orientation: II. Perception of the Upright With Displaced Visual Fields and With Body Tilted. *Journal of Experimental Psychology*, 38, 455–477.
- Attneave, F. (1955). Symmetry, Information, and Memory for Patterns. *The American Journal of Psychology*, 68(2), 209–222.
- Baker, F. B. (1993). Sensitivity of the Linear Logistic Test Model to Misspecification of the Weight Matrix. *Applied Psychological Measurement*, 17(3), 201–210.
- Bell, H. H. & Handel, S. (1976). The role of pattern goodness in the reproduction of backward masked patterns. *Journal fo Experimental Psychology: Human Perception and Performance*, 2(1), 139-150.
- Beuscart-Zéphir, M.-C., & Beuscart, R. (1988). Tests de Performances: une Méthode d'Analyse des Stratégies de Résolution . *European Journal of Psychology of Education*, III(1), 33–51.
- Botzum, W. A. (1951). A Factorial Study of the Reasoning and Closure Factors. *Psychometrika*, 16(4), 361–386.
- Buffart, H. & Leeuwenberg, E. (1983). Structural Information Theory. In H. G. Geissler, E. L. J. Leeuwenberg, S. Link & V. Sarris (Eds.), *Modern issues in perception* (pp. 48–72). Berlin: Erlbaum.
- Buffart, H., Leeuwenberg, E. & Restle, F. (1981). Coding Theory of Visual Pattern Completion. *Journal of experimental psychology. Human perception and performance*, 7(5), 241–274.
- Bruce, V., Green, P.R. & Georgeson, M.A. (2004). *Visual perception: physiology, psychology, & ecology* (4th ed.). Hove: Psychology Press.
- Checkosky, S. F., & Whitlock, D. (1973). Effects of pattern goodness on recognition time in a memory search task. *Journal of experimental psychology*, 100(2), 341–848.
- Clement, D. E. & Varnadoe, K. W. (1967). Pattern uncertainty and the discrimination of visual patterns. *Perception & Psychophysics*, 2(9), 427-431.
- Dieterich, H. (1975). *Die Validität des kognitiven Stils der Feldabhängigkeit/ Feldunabhängigkeit für die pädagogische Diagnostik*. Diss., Universität Münster.

- Draxler, C. (2002). *Ein Vergleich mittels Simulationsstudien bezüglich der Trennschärfe des Likelihood-Ratio-Tests nach Andersen und dem MIRA-Ansatz von Rost*. Unveröff. Dipl. Arbeit, Universität Wien.
- Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah: Lawrence Erlbaum.
- Evans, C. Richardson, J. T. E., & Waring, M. (2013). Field independence: reviewing the evidence. *The British journal of educational psychology*, 83, 210–24.
- Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 37, 359—374.
- Fischer, G. H. (1974). Einführung in die Theorie psychologischer Tests. Bern: Huber.
- Fischer, G. H. (1995). The Linear Logistic Test Model. Fischer, G. H. & Molenaar, I. W. (Eds.), *Rasch Models: Foundations, Recent Developments, and Applications* (S. 131–155). New York: Springer.
- Fischer, G. H., & Formann, a. K. (1982). Some Applications of Logistic Latent Trait Models with Linear Constraints on the Parameters. *Applied Psychological Measurement*, 6(4), 397–416.
- Formann, A. K. (2008). *Skriptum zur Vorlesung Testtheorie und Testkonstruktion* (neue korrigierte Auflage). Wien: Facultas WUV.
- Gardner, R. W., Jackson, D. N. & Messick, S. J. (1960). Personality organization in cognitive controls and intellectual abilities. *Psychological Issues*, 2(4), gesamte Monographie 8.
- Garner, W. R. & Sutliff, D. (1974). The effects of goodness on encoding time. *Perception & Psychophysics*, 16(3), 426-430.
- Gittler, G. (1990). *Dreidimensionaler Würfeltest (3 DW)*. Weinheim: Beltz.
- Gittler, G. & Wild, B. (1989). Der Einsatz des LLTM zur Konstruktion eines adaptiven Tests. In K. D. Kubinger (Hrsg.), *Moderne Testtheorie - Ein Abriss samt neuesten Beiträgen* (S. 115-139). (2. verb. Aufl.) Weinheim [u.a.]: Beltz.
- Goldstein, E. B. (2002). *Wahrnehmungspsychologie* (2. dt. Aufl.). Heidelberg [u.a.]: Spektrum Akademischer Verlag.
- Goldstein, E. B. (2010). *Sensation and Perception* (8th ed.). Belmont: Wadsworth Cengage Learning.
- Goodenough, D. R., & Karp, S. A. (1961). Field dependence and intellectual functioning. *Journal of abnormal and social psychology*, 63(2), 241–6.
- Gordon, I. E. (2004). *Theories of Visual Perception* (3rd ed.). Hove: Psychology Press.
- Gottschaldt, K. (1929). *Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren: I. Die Wirkung gehäufter Einprägung von Figuren auf ihre Sichtbarkeit in umfassenden Konfigurationen*. Berlin: Springer.

- Hergovich, A. (1992). Computer-Häuschen-Test: die Erstellung einer Computerversion des Häuschen-Tests zur Messung der Feldabhängigkeit. Unveröff. Diss., Universität Wien.
- Hergovich, A. & Hörndl, H. (1994). *Gestaltwahrnehmungstest. Ein computerbasiertes Verfahren zur Messung der Feldartikulation. Test: Manual und Software*. Frankfurt am Main: Swets Test Services.
- Hofsten, C. von. & Spelke E. S. (1985). Object perception and object-directed reaching in infancy. *Journal of Experimental Psychology: General*, 114, 198–212.
- Hornke, L. F. & Rettig, K. (1989). Regelgeleitete Itemkonstruktion unter Zuhilfenahme kognitionspsychologischer Überlegungen. In K. D. Kubinger (Hrsg.), *Moderne Testtheorie - Ein Abriß samt neuesten Beiträgen* (S. 115-139). (2. verb. Aufl.) Weinheim [u.a.]: Beltz.
- Hulleman, J., & Humphreys, G. W. (2004). Is there an assignment of top and bottom during symmetry perception? *Perception*, 33(5), 615–620.
- King, D. B. & Wertheimer, M. (2005). *Max Wertheimer & Gestalt Theory*. New Brunswick, NJ [u.a.]: Transaction Publishers.
- Kogan, N. & Block, J. (1991). Field dependence- independence from early childhood through adolescence: Personality and socialization aspects. In S. Wapner & J. Demick (Eds.), *Field Dependence-Independence: cognitive style across the life span* (pp. 177-208). Hillsdale: Lawrence Erlbaum Associates.
- Kohs, S. C. (1920). The Block-design Tests. *Journal of Experimental Psychology*, 3(5), 357–376.
- Köhler, W. (1933) Psychologische Probleme. Berlin: Julius Springer.
- Koller, I, Alexandrowicz, R. & Hatzinger, R. (2012). Das Rasch-Modell in der Praxis: eine Einführung mit eRm. Wien: Facultas WUV.
- Köstlin-Gloge, G (1978). Kognitive Stile im Entwicklungsverlauf. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 10, 52-74.
- Kubinger, K. D. (1979). Das Problemlöseverhalten bei der statistischen Auswertung psychologischer Experimente. Ein Beispiel hochschuldidaktischer Forschung. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 26, 467- 495.
- Kubinger, K. D. (1989). Aktueller Stand und kritische Würdigung der Probabilistischen Testtheorie. In K. D. Kubinger (Hrsg.), *Moderne Testtheorie - Ein Abriß samt neuesten Beiträgen* (S. 19-83), (2. Aufl.). München: Psychologie Verlags Union.
- Kubinger, K. D. (2005). Psychological Test Calibration Using the Rasch Model – Some Critical Suggestions on Traditional Approaches. *International Journal of Testing*, 5(4), 377–394.
- Kubinger, K. (2008). On the revival of the Rasch model-based LLTM: From constructing tests using item generating rules to measuring item administration effects. *Psychology Science Quarterly*, 50(3), 311–327.

- Kubinger, K. D. (2009a). *Adaptives Intelligenz Diagnostikum - Version 2.2 (AID 2) samt AID 2-Türkisch*. Göttingen: Beltz.
- Kubinger, K. D. (2009b). *Psychologische Diagnostik: Theorie und Praxis psychologischen Diagnostizierens* (2. Aufl.). Göttingen: Hogrefe.
- Kubinger, K. D. & Hagenmüller, B. (in Arbeit). *Adaptives Intelligenz Diagnostikum 3 – Gruppe*. Göttingen: Hogrefe.
- Kubinger, K. D. & Holocher-Ertl, S. (in Druck). *Adaptives Intelligenz Diagnostikum 3*. Göttingen: Beltz.
- Kubinger, K. D. & Wurst, E. (1985). *Adaptives Intelligenz Diagnostikum (AID)*. Weinheim: Beltz.
- Kubinger, K. D. & Wurst, E. (2000). *Adaptives Intelligenz Diagnostikum - Version 2.1 (AID 2)*. Göttingen: Beltz.
- Kurian, G., Sharma, N. K., & Santhakumari, K. (1987). Cognitive strategies in Koh's Block Design test performance: A preliminary report. *Psychological Studies*, 32, 29-32.
- Leeuwenberg, E. L. J. (1969). Quantitative Specification of Information in Sequential Patterns. *Psychological Review*, 76(2), 216–220.
- Leeuwenberg, E. L. J. (1971). A Perceptual Coding Language for Visual and Auditory Patterns. *The American Journal of Psychology*, 84(3), 307–349.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Auflage). Weinheim: Beltz.
- Maderthaner, R. (2008). *Psychologie*. Wien: Facultas WUV.
- McKenna, F. P. (1984). Measures of field dependence: Cognitive style or cognitive ability? *Journal of Personality and Social Psychology*, 47(3), 593–603.
- Mair, P., Hatzinger, R. & Maier, M.J. (2013). Extended Rasch Modeling – Version 0.15-1 [online]. URL: <http://cran.r-project.org/web/packages/eRm/eRm.pdf> [01.10.2013]
- Nährer, W. (1977). Modellkontrollen bei Anwendung des linearen logistischen Modells in der Psychologie. Unveröff. Diss., Universität Wien.
- Nährer, W. (1980). Modellkontrollen bei der Anwendung des linearen logistischen Testmodells. *Diagnostica*, XXVI(2), 112—118.
- Needham, A. & Baillargeon, R. (1997). Object segregation in 8-month-old infants. *Cognition* 62(2), 121-149.
- Palmer, S. E. (1999). *Vision Science*. Cambridge: MIT Press.
- Palmer, S. E., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, 1(1), 29–55.

- Pemberton, C. (1952). The closure factors related to other cognitive processes. *Psychometrika*, 17(3), 267–288.
- Petermann, F. (2012). *Deutsche Bearbeitung der Wechsler Adult Intelligence Scale – Fourth Edition WAIS-IV*. Frankfurt: Pearson Assessment.
- Peterson, M. A. & Gibson, B. S. (1994). Object recognition contributions to figure-ground organization: operations on outlines and subjective contours. *Perception & psychophysics*, 56(5), 551–564.
- Peterson, M. A., & Salvagio, E. (2008). Inhibitory competition in figure-ground perception: Context and convexity. *Journal of Vision*, 8(16), 1–13.
- Poinstingl, H. (2009). The Linear Logistic Test Model (LLTM) as the methodological foundation of item generating rules for a new verbal reasoning test. *Psychology Science Quarterly*, 51(2), 123–134.
- Quinn, P. C., & Bhatt, R. S. (2001). Object recognition and object segregation in infancy: historical perspective, theoretical significance, “kinds” of knowledge, and relation to object categorization. *Journal of experimental child psychology*, 78(1), 25–34.
- R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org/>.
- Rasch, G. (1980). Probabilistic models for some intelligence and attainment tests (Reprint). Chicago: The University of Chicago Press.
- Restle, F. (1979). Coding Theory of the Perception of Motion Configurations. *Psychological Review*, 86(1), 1–24.
- Royer, F. L. (1977). Information processing in the block design task. *Intelligence*, 1, 32–50.
- Royer, F. L. & Weitzel, K. E. (1977). Effect of perceptual cohesiveness on pattern recoding in the block design task. *Perception & Psychophysics*, 21(1), 39–46.
- Rozencwajg, P. (1991). Analysis of Problem Solving Strategies on the Kohs Block Design Test. *European Journal of Psychology of Education*, VI(1), 73–88.
- Rozencwajg, P. & Corroyer, D. (2001). Strategy development in a block design task, 30, 1–25.
- Schorr, D., Bower, G. H. & Kiernan, R. (1982). Stimulus variables in the Block Design Task. *Journal of consulting and clinical psychology*, 50(4), 479–487.
- Sonnleitner, P. (2008). Using the LLTM to evaluate an item generating system for reading comprehension. *Psychology Science Quarterly*, 50, 345–362.
- Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, 14(1), 29–56.
- Spelke, E. S., Breinlinger, K., Jacobson, K. & Phillips, A. (1993). Gestalt relations and object perception: a developmental study. *Perception*, 22(12), 483–501.

- Spelke, E. S., Hofsten, C. von. & Kestenbaum, R. (1989). Object perception in infancy: Interaction of spatial and kinetic information for object boundaries. *Developmental Psychology, 25*(2), 185–196.
- Sternberg, R. J., & Grigorenko, E. L. (1997). Are cognitive styles still in style? *American Psychologist, 52*(7), 700–712.
- Thurstone, L. L. (1938). The perceptual factor. *Psychometrika, 3*(1), 1–17.
- Thurstone, L. L. (1944). *A factorial study of perception*. Chicago: University of Chicago Press.
- Thurstone, L. L. (1949). *Mechanical aptitude III: analysis of group tests*. Psychometric Laboratory Reports, 55. University of Chicago.
- Van Meel, J. (1991) Proximate social influences on field dependence-independence. In S. Wapner & J. Demick (Eds.), *Field Dependence-Independence: cognitive style across the life span* (pp. 325–251). Hillsdale: Lawrence Erlbaum Associates.
- Vecera, S. P., & Palmer, S. E. (2006). Grounding the figure: surface attachment influences figure-ground organization. *Psychonomic Bulletin & Review, 13*(4), 563–569.
- Vecera, S. P., Vogel, E. K., & Woodman, G. F. (2002). Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General, 131*(2), 194–205.
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M. & Heydt, R. von der. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*(6), 1172–1217.
- Waldenmair, M. (1984). *Zur Dimensionalität des Feldabhängigkeitskonzepts*. Unveröff. Diss., Universität Wien.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition*. San Antonio: Harcourt
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II. [Investigations in Gestalt Theory: II. Laws of organization in perceptual forms]. *Psychologische Forschung, 4*, 301–350.
- Witkin, H. A. (1976). Cognitive style in academic performance and teacher-student relations. In S. Messick (Ed.), *Individuality in Learning: Implications of Cognitive Style and Creativity for Human Development*. San Francisco: Jossey-Bass.
- Witkin, H. A. & Asch, S. E. (1948a). Studies in Space Orientation: III. Perception of the Upright in the Absence of a Visual Field. *Journal of Experimental Psychology, 38*, 603–614.
- Witkin, H. A. & Asch, S. E. (1948b). Studies in Space Orientation: IV. Further Experiments on Perception of the Upright with Displaced Visual Fields. *Journal of Experimental Psychology, 38*, 762–782.
- Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R. & Karp, S. A. (1962). *Psychological Differentiation. Studies of Development*. Oxford: Wiley.

- Witkin, H. A., Goodenough, D. R., & Karp, S. A. (1967). Stability of cognitive style from childhood to young adulthood. *Journal of personality and social psychology*, 7(3), 291–300.
- Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B. & Wapner, S. (1972). *Personality through perception* (Nachdruck). Westport: Greenwood Press.
- Witkin, H. A., Oltman, P. K., Raskin, E. & Karp, S. A. (1971). *A Manual for the Embedded Figures Test*. Palo Alto: Consulting Psychologists Press.
- Young, H. H. (1959). A test of Witkin's field-dependence hypothesis. *Journal of Abnormal and Social Psychology*, 59(2), 188–192.

# X. Anhang

---

## 1. Testperson-Notizen (TPN)

Es handelt sich um die strukturierte und ausformulierte Zusammenfassung handschriftlicher, stichwortartiger Notizen, die während und nach der Testbearbeitung verfasst wurden. Die Schreibweise sowie die Verwendung bestimmter Begriffe und Abkürzungen ist willkürlich, es liegt ihnen keine explizite Definition oder Rechtfertigung zugrunde.

Abkürzungen:

TPN ..... Testperson-Notiz

**Tabelle 18:** Testperson-Notizen (TPN) über introspektive Beobachtungen, auffällige Merkmale der Materialeigenschaften sowie darauf begründete Annahmen zum Lösungsprozess bei der Bearbeitung des UT 10 in der Position als Testperson

<b>Testperson-Notizen</b>	
<b>Vorlagekärtchen</b>	
TPN_01	Die Vorlage für das Muster ist auf einem Kärtchen x*x cm abgebildet.
TPN_02	Das Kärtchen an sich ist weiß. Die Vorlage ist weiß und rot und durch einen schwarzen Rahmen abgegrenzt. Das Weiß des Kärtchens und das Weiß der Vorlage sind ident. Der schwarze Rahmen definiert, was als weißes Musterelement zur Vorlage gehört und was nicht berücksichtigt werden muss.
TPN_03	Die Größe der Vorlage beträgt zwischen ..... cm. Die Vorlage ist kleiner als das mit den Würfeln nachgelegte Muster.
<b>Muster</b>	
TPN_04	Wenn alle Items gleichzeitig betrachtet werden, so wird eine Systematik erkennbar, welche sich nicht sofort in Worte fassen lässt. Es scheinen jedenfalls alle Muster konstruiert und nicht zufällig zusammengewürfelt.

<i>TPN_05</i>	Ein sich aufdrängendes Merkmal vieler Muster ist deren <b>Symmetrie oder Regelmäßigkeit</b> , welche eine gewisse Beziehung zur zugrundeliegenden Struktur aufzuweisen scheinen.
<i>TPN_06</i>	Items <i>14</i> , <i>15</i> und <i>18</i> scheinen hier am weitesten von der eher einheitlichen <i>Charakteristik</i> der Aufgaben abzuweichen.
<i>TPN_07</i>	Besonders auffällig scheint auch Item <i>6.1</i> , bei welchem ich nicht umhin kann, die rote Fläche als Häuschen mit Garage zu identifizieren, während bei allen anderen Items kein solches Erkennen von bekannten Objekten gelingt.
<i>TPN_08</i>	Die Items <i>3</i> , <i>8</i> , <i>5.1</i> und <i>16</i> lassen sich allein durch die diagonal geteilten Würfelseiten nachlegen.
<i>TPN_09</i>	Nur ein einziges Item, nämlich <i>5</i> , besteht aus lediglich 2 Würfeln und es ist damit gleichzeitig sowohl das einzige, das nicht quadratisch ist, als auch das einzige, das nur aus <i>vollen</i> Würfelseiten zusammengesetzt wird.
<i>TPN_10</i>	7 der Muster setzen sich aus 4 im Quadrat zueinanderstehenden Würfeln zusammen und die übrigen 10 Muster lassen sich durch 9 quadratisch angeordnete Würfel nachlegen.

#### ***Strukturierung***

<i>TPN_11</i>	Das Wissen, wie viele Würfel erforderlich sind, ist sehr hilfreich bei der Strukturierung, da in der Anzahl der Würfel kombiniert mit der Form der Vorlage (meist Quadrat) im Grunde die gesamte Information vorhanden ist um das Muster in die entsprechenden Würfelflächen zu gliedern. Bei denselben Kombinationen sind immer dieselben <i>Raster</i> anzuwenden. Die Anzahl der Würfel bestimmt, ob halbiert oder gedrittelt werden muss, und es bleibt lediglich das Einschätzen der Relationen als Herausforderung.
<i>TPN_12</i>	Die Relationen der Raster können leichter eingeschätzt werden, wenn der Figurumriss einen Rückschluss auf die Würfelkanten erlaubt – sowohl horizontal als auch vertikal. Am informativsten scheinen Hinweise auf die Eckpunkte von Würfelflächen. Z.B. weisen die diagonal geteilten Würfelflächen in den Spitzen auf einen relevanten Eckpunkt hin, die gerade halbierten Würfelflächen leiten den Blick eher in die <b>Irre</b> [s8].

### ***Strategie und Vergleich***

<i>TPN_13</i>	Da die Vorlage während der Reproduktion sichtbar liegen bleibt kann beliebig oft der eigene Fortschritt kontrolliert werden. Es muss einerseits das Zielmuster nicht notwendigerweise gemerkt werden und andererseits erlaubt es auch ein stärker synthetisches denn analytisch-strukturierendes Vorgehen gleichermaßen, da durch Versuch und Irrtum beliebige Würfelkombinationen auf ihre Passung mit dem Vorlagemuster innerhalb des Blickfeldes kontrolliert werden können.
<i>TPN_14</i>	Letzteres ist ein mögliches, aber wohl langsameres Vorgehen, das wohl eher bei Aufgaben angewandt wird, die einem schwerer fallen. Um die Aufgabe innerhalb der Zeitvorgabe zu lösen scheint daher ein höheres Mindestmaß an Strukturierungsleistung vorausgesetzt. Es scheint schneller, die nötige Würfeloberfläche aus der Vorlage abzuleiten als den Würfel lange an der Position zu drehen. Dafür muss die Würfelfläche im Muster als vom Rest unabhängige Einheit erkannt und identifiziert werden.
<i>TPN_15</i>	Symmetrische oder andere leicht zu merkende Figuren brauchen weniger Vergleiche mit der Vorlage, als unregelmäßige Formen.
<i>TPN_16</i>	Die notwendige Würfeloberfläche muss beim Drehen auch dann als solche erkannt werden, wenn sie nicht in exakt derselben Position auftaucht. Dies betrifft die drei „nicht vollen“ Würfeloberflächen. Auch wenn die tatsächliche Rotation dann mit den Händen geschieht, so muss zumindest soweit „mental rotiert“ werden, als dass eine Seite als jene, die hingelegt werden muss, identifiziert werden kann – und nicht einfach zur nächsten Seite weitergedreht wird. Auch wenn danach an der Position drehend der Vergleich mit der Vorlage gesucht wird.

## 2. Testleiter-Notizen (TLN)

Es handelt sich um die strukturierte und ausformulierte Zusammenfassung handschriftlicher, stichwortartiger Notizen, die während und nach der Testbearbeitung verfasst wurden. Die Schreibweise sowie die Verwendung bestimmter Begriffe und Abkürzungen ist willkürlich, es liegt ihnen keine explizite Definition oder Rechtfertigung zugrunde.

Abkürzungen:

TLN ..... Testleiter-Notiz  
TL ..... Testleiter  
TPn ..... Testperson  
UT10 ..... Untertest *Analysieren und Synthetisieren* im AID 3  
UT10a ..... Untertest *Strukturieren* im AID 3

**Tabelle 19:** Testleiter-Notizen (TLN) über Gelegenheitsbeobachtungen und darauf begründete Annahmen zum Lösungsprozess bei der Vorgabe der Untertests 10 und 10a

<b>Testleiter-Notizen</b>	
<b>Eindrücke zur Fähigkeitsverteilung</b>	
TLN_01	TL haben oft das Gefühl, <b>das Kind kann es eben oder nicht</b> – der mittlere Fähigkeitsbereich wird seltener wahrgenommen.
TLN_02	Die Leistung im UT10 scheint stark mit der im UT10a zusammenzuhängen.
<b>Beobachtungen zu Strategie und Vorgehen</b>	
TLN_03	Die Kinder <b>beginnen das Nachlegen meist an einer Ecke</b> und gehen dann ganz unterschiedlich weiter vor.  Beispielsweise werden dann entweder zunächst die anderen Ecken gelegt, es wird Reihe für Reihe (sowohl von links nach rechts als auch von oben nach unten) vorgegangen oder es werden von einer Ecke ausgehend die nächsten immer an das vorhandene angelegt usw..

TLN_04	<p>Auch Viellöser machen häufig Fehler beim Instruktionsitem. Hierbei scheint meist der Bereich, der das Muster bestimmt, nicht eindeutig erkannt zu werden.</p> <p>→ beim roten: Oft bilden die Kinder die rote Fläche durch 4 „viertel“- Würfel. Dies legt zwei Schlüsse nahe:</p> <ol style="list-style-type: none"> <li>1. Die Kärtchenfläche, die nicht zur Vorlage gehört, wird eventuell mit einbezogen.</li> <li>2. Die Übertragungsleistung von der kleineren Vorlage auf die größeren Würfelflächen ist womöglich nicht implizit anzunehmen.</li> </ol>
--------	---

#### ***Beobachtungen bei Viellösern***

TLN_05	<p>Sehr erfolgreiche TPn im <b>UT10</b> scheinen „auf den ersten Blick“ zu erkennen, welche Oberflächen notwendig sind – sie gehen seltener Würfel für Würfel vor und drehen nicht jeden einzelnen, sondern sortieren von den daliegenden gleich die aus, die sie mit der jeweiligen obenliegenden Fläche brauchen können und legen sie oft auch schon an die entsprechende Position, unabhängig davon, ob die Würfel damit allein stehen.</p> <p><i>typisch:</i> zusammenfassende Bewegung in der erst am Schluss die Würfel, die nicht gleich angrenzend hingelegt wurden, zum Rest geschoben werden und damit das fertige Muster ohne Unterbrechung daliegt</p>
TLN_06	<p>Sehr erfolgreiche TPn im <b>UT10a</b> zeichnen gleich das Raster mit durchgängigen Strichen ein (Orientierung an den Kanten – falls vorhanden – besonders hilfreich).</p> <p>Erfolgreiche TPn im <b>UT10a</b> gehen Würfel für Würfel vor (Orientierung scheinbar an den möglichen Würfelloberflächen) und kommen so zur Lösung, brauchen aber länger als TPn, die Raster ziehen.</p> <p>Die beiden Vorgehensweisen schließen sich nicht gegenseitig aus, sondern scheinen sich ergänzen zu können. Manche TPn</p>

---

### ***Beobachtungen bei Weniglösern***

---

<i>TLN_07</i>	<p>Sehr schwache TPn im <b>UT10</b> gehen meist Würfel für Würfel vor und drehen den Würfel nicht nur in der Hand sondern oft auch mehrmals auf dem Tisch am anderen Würfel anliegend, bis dass eine Übereinstimmung mit der Vorlage erkannt wird.</p> <p><i>typisch:</i> ständig zwischen Vorlage und Würfel pendelnder Blick</p>
<i>TLN_08</i>	<p>Wenig erfolgreiche TPn im <b>UT10a</b> zeichnen oft konsequent dasselbe Raster mit nur geringfügiger Anpassung ein oder ziehen die Linien entlang der Feld- bzw. Figurumrisse.</p> <p>Bei wenig erfolgreichen TPn ist in der Regel eine langsamere Bearbeitung zu beobachten.</p>

---



### **3. Lebenslauf**

#### **Sabrina Köb**

---

geboren am  
Staatsbürgerschaft

[REDACTED]  
österreichisch

#### **Ausbildung**

---

seit 2007

**Studium der Psychologie**, Universität Wien

Schwerpunkte:

1. Psychologische Diagnostik  
Eignungs- und Leistungsdiagnostik, Teilleistungsstörungsdiagnostik, psychologisch-diagnostische Verfahren
2. Wirtschaftspsychologie  
Organisationssberatung (Führung, Change Management), betriebliches Gesundheitsmanagement

2007 – 2011

**Studium der Publizistik- und Kommunikationswissenschaften**,  
Universität Wien, ohne Abschluss

2001 – 2006

**Höhere Lehranstalt für wirtschaftliche Berufe**, Rankweil  
mit gutem Erfolg abgeschlossen  
Reife- und Diplomprüfung mit gutem Erfolg bestanden  
Schwerpunkt:  
dritte lebende Fremdsprache (Spanisch – Castellano)

1997 – 2001

**Hauptschule Strohdorf**, Wolfurt

1993 – 1997

**Volksschule Bütze**, Wolfurt

#### **Berufserfahrung und Praktika**

---

05/ 2013 – 12/2013

**Service mit Inkasso** – Caffè Latte, Wien (geringfügig)

10/ 2011 – 12/2011

**Personalberatung** – Hill Woltron Management Partner, Wien  
Praktikum (Teilzeit)

Tätigkeiten:

- ✓ Suche und Identifikation geeigneter Bewerber/innen in Abhängigkeit des jeweiligen Anforderungsprofils
- ✓ Mithilfe bei Direktansprache
- ✓ Bewertung und Vorselektion schriftlicher Bewerbungsunterlagen
- ✓ Protokollierung bei Interviews
- ✓ Vorgabe der *Hill*-Testbatterie
- ✓ Erstellen von Gutachten und Kandidatenberichten in deutscher und englischer Sprache
- ✓ Zusammenstellung von Präsentationsunterlagen für Kunden
- ✓ administrative Tätigkeiten

Sommermonate 2005 – 2011

**Service mit Inkasso** – diverse Saisonanstellungen, Vorarlberg

02/2007 – 06/2007

**Service mit Inkasso** – Hotel Krone, Dornbirn

11/2006 – 12/2006

**Au-Pair** – Madrid, Spanien

07/2004 – 09/2004

**Service Praktikum** – Dunbrody Country House Hotel & Restaurant, Irland