

Hager, Willi; Hübner, Sigrid

## **Denkförderung und Strategieverhalten: vergleichende Evaluation zweier Denkförderprogramme**

*Praxis der Kinderpsychologie und Kinderpsychiatrie* 47 (1998) 5, S. 277-301

urn:nbn:de:0111-opus-8969

Erstveröffentlichung bei:



[www.v-r.de](http://www.v-r.de)

### **Nutzungsbedingungen**

pedocs gewährt ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit dem Gebrauch von pedocs und der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### **Kontakt:**

**peDOCS**

Deutsches Institut für Internationale Pädagogische Forschung (DIPF)

Informationszentrum (IZ) Bildung

Schloßstr. 29, D-60486 Frankfurt am Main

eMail: [pedocs@dipf.de](mailto:pedocs@dipf.de)

Internet: [www.pedocs.de](http://www.pedocs.de)

# Praxis der Kinderpsychologie und Kinderpsychiatrie

Ergebnisse aus Psychoanalyse,  
Psychologie und Familientherapie

47. Jahrgang 1998

---

## Herausgeberinnen und Herausgeber

Manfred Cierpka, Heidelberg – Gunther Klosinski, Tübingen –  
Ulrike Lehmkuhl, Berlin – Inge Seiffge-Krenke, Mainz –  
Friedrich Specht, Göttingen – Annette Streeck-Fischer, Göttingen

## Verantwortliche Herausgeberinnen

Ulrike Lehmkuhl, Berlin  
Annette Streeck-Fischer, Göttingen

## Redakteur

Günter Presting, Göttingen

---

**V&R** Verlag Vandenhoeck & Ruprecht in Göttingen

## Denkförderung und Strategieverhalten: vergleichende Evaluation zweier Denkförderprogramme<sup>1</sup>

Willi Hager und Sigrid Hübner

### Summary

#### Cognitive Training and Thinking Strategies: a Comparative Evaluation

Combining a non-comparative with a comparative evaluation, two modern programs for fostering inductive reasoning, namely the German version of the "Cognitive training for children" by KLAUER and PHYE (1994; KLAUER 1989), and the "DenkMit" by SYDOW and MEINCKE (1994), are compared to each other and to a control program which intends to enhance aspects of memory instead of inductive reasoning. The programs were performed with N=49 children between six and eight years who had been postponed from regular school because of various reasons or who had been selected as especially in need for particular interventions from first classes. Besides the psychometric tests often used for assessing inductive reasoning, i.e. three subtests of the German form of the Culture Fair Test by CATTELL (WEISS a. OSTERLAND 1980), tasks of concept formation were applied for assessing changes in strategic behavior of children – a type of task which has been used in connection with inductive reasoning since many years. Counter to expectations, the children whose memory was trained, showed changes in performance in the psychometric test in a similar size as the children whose inductive reasoning was trained. These effects are interpreted in terms of special attention directed to the children during the intervention situation. Moreover, it was found that despite the authors claim to the opposite the DenkMit did not cause any changes in visual perception. In contrast to the author's intentions, the "Cognitive Training for Children" did cause some substantive changes in the area of visual perception. The pattern of results with the concept formation tasks, however, overall indicates that the reasoning programs caused some changes in strategic behaviors of the children. Although these changes are not very impressive, they cannot be attributed to extraneous factors such as special attention.

<sup>1</sup> Die Studie entstand im Rahmen eines von der DFG geförderten Forschungsprojektes (HA 2502/1-1).

## Zusammenfassung

In einer Kombination von vergleichender und isolierter (nicht-vergleichender) Evaluation werden zwei Denkförderprogramme als Alternativprogramme, nämlich das Denktraining für Kinder I von KLAUER (1989) und das DenkMit von SYDOW und MEINCKE (1994), mit einem Gedächtnisförderprogramm als Kontrollprogramm verglichen. Die Intervention erfolgte mit N=49 sechs- bis achtjährigen Kindern, die überwiegend für ein Jahr von der regulären Einschulung zurückgestellt oder die von den Lehrerinnen zweier erster Klassen als schwach und daher besonders förderungsbedürftig ausgewählt worden waren. Neben in einschlägigen Evaluationen zu den Denkprogrammen üblicherweise eingesetzten psychometrischen abhängigen Variablen (Subtests aus dem CFT 1 von WEISS u. OSTERLAND 1980) werden selbst konstruierte Konzeptbildungsaufgaben eingesetzt, also ein Aufgabentyp, der seit langer Zeit in der Kognitionspsychologie zur Erfassung des Strategieverhaltens beim induktiven Denken verwendet wird; Strategieverbesserungen beim induktiven Denken stellen bei den Denkprogrammen zentrale Ziele dar. Beim psychometrischen Test zur Erfassung induktiver Anforderungen stellender Denkleistungen ergaben sich erwartungswidrig auch unter dem Kontrollprogramm Veränderungen in der gleichen Größenordnung wie unter den Denkförderprogrammen, die für das Wirksamwerden von Zuwendungswirkungen sprechen. Des weiteren kann der Anspruch des DenkMit der Förderung auch im Wahrnehmungsbereich nicht bestätigt werden, während das Denktraining entgegen dem Anspruch des Programmautors gerade in diesem Bereich zu substantiellen Verbesserungen führt. Die Wahrnehmungsleistungen werden mit der POD von SAUTER (1979) erfaßt. Für die anhand der Konzeptbildungsaufgaben erhobenen abhängigen Variablen (AVn) ergeben sich differentielle Wirkungen. Das Befundmuster spricht insgesamt für eine wenn auch schwache Wirkung der beiden Denkförderprogramme im Strategiebereich, die sich vielleicht durch eine entsprechende Änderung besonders am Denktraining I verstärken ließe.

### 1 Untersuchte Förderprogramme und Anliegen des Versuchs

Um für die praktische Arbeit von Nutzen sein zu können, muß für jede Art von Fördermaßnahme nachgewiesen werden, daß sie eine Verbesserung in dem angestrebten Bereich bewirkt (Wirksamkeit), daß diese Verbesserung nicht nur kurzfristig ist und daß ein Transfer auf ähnliche Aufgaben und Probleme bewirkt wird. Die grundsätzliche Wirksamkeit eines Programms wird dabei im Vergleich zu einer Kontrollgruppe ermittelt, die ebenfalls einer Intervention unterzogen wird, die aber nicht auf die Bereiche des zu evaluierenden Programmes abzielt (Kontrollprogramm; isolierte oder nicht-vergleichende Evaluation). Dies soll gewährleisten, daß die Rahmenbedingungen der Wirksamkeitserprobung oder der Evaluation (annähernd) gleich sind, so daß festgestellte Unterschiede mit hinreichender Sicherheit den Inhalten des zu evaluierenden Programmes zugeschrieben werden können; mit sogenannten Wartegruppen, die keiner Intervention unterzogen werden, kann dieses Ziel in aller Regel nicht erreicht werden. Neben der Feststellung der grundsätzlichen Wirksamkeit eines Förderprogrammes interessiert zudem häufig, ob zwei (oder mehr) Programme, die das gleiche Interventionsziel aufweisen, dies aber mit verschiedenen Vor-

gehensweisen zu erreichen versuchen (Alternativprogramme; vergleichende Evaluation), sich hinsichtlich ihrer vergleichenden Wirksamkeit voneinander unterscheiden. Die Ergebnisse derartiger Studien können als Grundlage für die Beurteilung der Effizienz solcher Alternativprogramme dienen, wobei in die Effizienzbeurteilung noch zusätzliche Gesichtspunkte eingehen, wie etwa materielle und immaterielle Kosten bei der Beschaffung und Durchführung der Programme sowie ihre „Robustheit“ in unterschiedlichen Fördersituationen und bei unterschiedlichen fördernden Personen. Derartige Untersuchungen lassen zwar stets vergleichende Aussagen zu, aber es bleibt unter Umständen offen, ob beide Programme gleich wirksam oder gleich unwirksam sind. Deshalb ist es günstig, in einer Untersuchung beide Formen der Evaluation miteinander zu verbinden, da auf diese Weise auch feststellbar ist, ob die Programme in der gewählten Fördersituation überhaupt wirksam sind (Vergleich beider Alternativprogramme mit dem Kontrollprogramm).

In den letzten Jahren sind mehrere Denkförderprogramme entstanden. Wir befassen uns in dieser Arbeit mit dem Denktraining für Kinder I von KLAUER (1989a) und mit dem DenkMit von SYDOW und MEINCKE (1994). Mit dem *Denktraining*, das wesentlich auf Überlegungen des Programmautors zum induktiven Denken beruht (KLAUER 1989a) und das für einen Altersbereich von fünf bis sieben Jahren vorgesehen ist (nach KLAUER 1995a ist es für Sechs- bis Achtjährige gedacht), soll das induktive Denken gefördert werden, das für KLAUER (1993a, S. 17) „in der Entdeckung von Regelmäßigkeiten durch Feststellung der Gleichheit oder der Verschiedenheit oder der Gleichheit *und* Verschiedenheit bei Merkmalen und Relationen bei ... (verschiedenartigem) Material“ besteht. Diese Aussage stellt zwar für KLAUER (1989a, S. 16; 1991, S. 15) eine „Theorie“ dar, aber sie verbietet keine weiteren Aussagen, sie ist als Aussage weder verifizierbar noch falsifizierbar, und sie kann daher keineswegs als eine Theorie fungieren (vgl. auch MERZ 1992; PATRY 1994). Aber diese Einschätzung ist für das Programm selbst völlig unproblematisch, denn dieses muß vor allem nachweisbar wirksam oder *effektiv* sein (HERRMANN 1984), und die Effektivität ist gänzlich unabhängig von der Frage der Güte der theoretischen Fundierung des Programmes (vgl. BRANDTSTÄDTER 1982, S. 27-30).

Wir sprechen daher – im übrigen auch mit KLAUER (1989a, S. 17) – etwas weniger anspruchsvoll von einem „Definitionssatz“, und dieser legt nahe, daß zwei distinkte Prozesse gefördert werden sollen. Die „Feststellung der Gleichheit und Verschiedenheit“ ist ein Prozeß, der üblicherweise mit dem Namen „Wahrnehmungsgeschwindigkeit“ belegt wird (vgl. etwa KAIL u. PELLEGRINO 1989, S. 30). Ob man dann im Anschluß an diesen Prozeß eine weitere kognitive Operation ausführt, mittels derer aus den identifizierten Gemeinsamkeiten eine Regel abgeleitet wird, ist u.E. eine Frage der entsprechenden *Instruktion* oder Aufforderung – die Regelableitung erfolgt (zumindest meist) *nicht* automatisch (vgl. auch HAGER u. HASSELHORN 1995a). Dennoch kann dies natürlich durch geeignete Instruktionen und Aufgaben in einem Programm durch Übung grundsätzlich erreicht werden.

Ungeachtet dessen hat die Hierarchie der angesprochenen kognitiven Prozesse HAGER und HASSELHORN (1993b, 1995b) dazu veranlaßt, die Wirkungen des Denktrainings I in einen Wahrnehmungsteil und einen darauf aufbauenden Regelableitungsteil aufzugliedern. Da das Erkennen von Gemeinsamkeiten und Unterschieden gezielt geübt wird, sollte nach unseren Vorstellungen auch das Denktraining I zu Verbesserungen im Wahrnehmungsbereich führen. Nicht

gänzlich, aber doch weitgehend unabhängig von derartigen Verbesserungen sind zusätzlich Verbesserungen im Bereich des Aufstellens von Regelhaftigkeiten (induktives Denken) zu erwarten.

Allerdings gibt es derzeit auf dem deutschsprachigen Markt immer noch kein standardisiertes Verfahren, das explizit mit der Aufforderung zu Regelableitung arbeitet und das aufgrund dessen geeignet wäre, das so verstandene induktive Denken zu erfassen (HAGER u. ELSNER 1995; HAGER u. HASSELHORN 1995a); dies gilt für den von BÜCHEL und SCHARNHORST (1993) vorgestellten Test ebenso wie für den von PERLETH in seiner Dissertation (1992) eingesetzten Mengen-Folge-Test, nicht jedoch für den „Test des induktiven Denkens“ (TID) von Hager et al. (1994). Aber diesem Test ist genau wie allen anderen gemeinsam, daß die induktive Anforderungen stellenden Aufgaben nicht unabhängig von der visuellen Wahrnehmung sind: Induktives Denken ist in der gleichen Weise wie viele andere kognitive Leistungen – zumindest im Kindesalter – nicht unabhängig von der visuellen Wahrnehmung erfaßbar.

KLAUER (1989a) argumentiert nun im wesentlichen, daß erstens der Faktor „Induction“ oder „Inductive Reasoning“ sehr hoch auf der allgemeinen fluiden Intelligenz ( $g_f$ ) i.S. der Unterscheidung CATTELLS (vgl. HORN u. CATTELL 1966) lade und daß zweitens die Mehrheit der Aufgaben in einschlägigen Intelligenztests wie dem Cattellschen CFT (WEISS u. OSTERLAND 1980) und den Ravenschen Matrizen (CPM; BECKER et al. 1980) über induktive Denkprozesse zu lösen seien. Dies rechtfertige es, etwa die Subtests 3, 4 und 5 des CFT 1 und auch die CPM als Tests zur Erfassung des induktiven Denkens anzusehen (vgl. auch KLAUER 1989b, 1995b, 1995c). In früheren Untersuchungen zu den CPM hat sich jedoch gezeigt, daß die Mehrheit der 36 Aufgaben „rein wahrnehmungsmäßig“ zu lösen ist (BECKER et al. 1989; HAGER u. ELSNER 1995), also letztlich mit geringerem kognitiven Aufwand, als es das zusätzlich zu aktivierende induktive Denken darstellt.

Für den CFT 1 ergibt sich zudem die Schwierigkeit, daß der Subtest 4 („Ähnlichkeiten“) in der gleichen Form etwa in „Frostigs Entwicklungstest der visuellen Wahrnehmung“ (FEW; LOCKOWANDT 1993) als auch in der „Prüfung der optischen Differenzierungsfähigkeit“ von SAUTER (1979; POD) auftaucht – allerdings jeweils nicht zur Erfassung des induktiven Denkens, sondern als Wahrnehmungs(sub)test. In dem Konzentrationstraining von KROWATSCHKE (1994) wird der gleiche Typ von Aufgabe zur Erhöhung der Konzentrationsleistung eingesetzt (Aufgabentyp 5.6: „Gleiche Figuren finden“; S. 124, S. 141; Aufgabentyp „Gleiches Muster finden“; S. 144, S. 149). Dies alles erschwert es, den besagten Subtest 4 aus dem CFT 1 („Ähnlichkeiten“) mit dem induktiven Denken im Sinne einer Regelableitung in Verbindung zu bringen. Und selbst die Subtests 3 und 5 des CFT 1 sind nicht mit einer Instruktion versehen, die auf das Erkennen einer Regelhaftigkeit hinausläuft.

Trotz dieser möglichen Probleme setzen wir die Subtests 3, 4 und 5 (Summe 3 oder  $\Sigma 3$ ) des CFT 1 ein, da dies erstens vom Programmautor sozusagen zur Erfassung „seines“ induktiven Denkens vorgegeben wird (KLAUER 1989b) und da zweitens diese drei Subtests bzw. die  $\Sigma 3$  eine wichtige Möglichkeit darstellt, die eigenen Ergebnisse mit denen in vorausgehenden Evaluationen zu vergleichen.

Mit den Subtests 1 und 2 (Summe 2 oder  $\Sigma 2$ ) des CFT soll laut Testautoren vor allem die Wahrnehmungsgeschwindigkeit, der Wahrnehmungsumfang und eine visuomotorische Komponente erfaßt werden (WEISS u. OSTERLAND 1980, S. 19). KLAUER führt jedoch aus: „Wenn man die Wahrnehmungsgeschwindigkeit bei Kindern erfassen will, so sind die Subtests 1 und 2 des CFT 1 keinesfalls eine gute Wahl“ (1995b, S. 298). Um Wahrnehmungsaspekte zu erfassen, haben wir daher die „Prüfung optischer Differenzierungsleistungen“ (POD) von SAUTER (1979) eingesetzt (zu den ebenfalls benutzten Subtests 1 und 2 des CFT 1 siehe HAGER u. HÜBNER 1996).

Die Aufgaben dieses Tests verlangen von den Kindern das Vergleichen von Merkmalen sowie das Erkennen von Gemeinsamkeiten und Verschiedenheiten. Man kann also wohl feststellen, daß dieser Test eher geeignet ist, mit dem Denktraining 1 auf der Wahrnehmungsseite Geübtes zumindest teilweise zu erfassen. Mithilfe der abhängigen Variable (AV) POD wird es also möglich, die Hypothese zu prüfen, daß mit dem Denktraining (auch oder vor allem) eine Verbesserung im Wahrnehmungsbereich erzielt wird („Wahrnehmungshypothese“ von HAGER u. HASSELHORN 1993b, 1995).

SYDOW und MEINCKE (1994) haben zwei Denkförderprogramme entwickelt, und zwar eines für den Altersbereich von drei bis vier Jahren und ein weiteres für den Bereich von fünf bis sechs Jahren. Wir befassen uns hier nur mit dem Programm für die älteren Kinder, für die auch das Denktraining von KLAUER in Frage kommt. Die Autoren gehen in ihrer Handanweisung zum Programm *DenkMit für 5- und 6jährige* zwar nicht explizit auf die theoretische Grundlage ihres Programms ein, aber es dürfte wesentlich auf der Intelligenzkonzeption von KLIX (z.B. 1983) beruhen, der einen engen Zusammenhang zwischen Denk- und Wahrnehmungsleistungen annimmt. Demzufolge streben die Programmautoren auch eine Förderung des Denkens und der Wahrnehmung an. Ihr Programm für die älteren Kinder (s.o.) umfaßt mehrheitlich Analogieaufgaben, die in acht zunehmend schwieriger werdenden Aufgabengruppen zusammengefaßt und die mittels Puzzles zu lösen sind. Zur Erfassung der Programmwirksamkeit setzen SYDOW und MEINCKE (1994) vor allem (allerdings nicht ausschließlich; siehe S. 70-71) einen selbstkonstruierten Test ein, dessen Aufgaben sich mit bestimmten Aufgaben im Übungsprogramm decken, so daß also „eine programmnahe Erfassung der Wirksamkeit“ erfolgt. Diese erschwert es aber, reine Coachingwirkungen („Etwas wird intensiv geübt und nach der Übung auf die gleichen Aufgaben angewendet“; vgl. HAGER u. ELSNER 1996; HASSELHORN 1995; MERZ 1992) von darüber hinausgehenden Programmwirkungen und seinem *Transfer* zu trennen. Für die Ergebnisse in einem trainingsfernen Test erfolgt keine signifikanzstatistische Absicherung, so daß nicht beurteilt werden kann, ob die berichteten Zuwächse noch durch den Zufall erklärbar sind oder ob sie „echte“ Trainingswirkungen darstellen (SYDOW u. MEINCKE 1994, S. 71). In einer jüngeren Untersuchung, in der das DenkMit mit dem Denktraining von KLAUER (1989a) verglichen wurde (SCHMUDE u. SYDOW 1994), ist jedoch der gesamte CFT 1 verwendet worden, der zunächst als angemessen in bezug auf die beiden Programmziele „Denken und Wahrnehmen“ angesehen werden kann. Da das DenkMit jedoch überwiegend Analogieaufgaben enthält, die als paradigmatisch für das *induktive* Denken gelten (siehe z.B. PELLEGRINO u. GLASER 1980; RESING 1996; STERNBERG u. GARDNER 1983), darf vermutet werden, daß auch das von ihm fokussierte Denken das *induktive* Denken ist. Diese Vermutung wird durch den Umstand erhärtet, daß die Subtests 3, 4 und 5 des CFT 1 von KLAUER (1989b) zu Erfassung des induktiven Denkens als geeignet angesehen werden (s.o.). Zusätzlich sind SCHMUDE und SYDOW (1994) offenbar davon ausgegangen, daß mit den beiden „nicht-induktiven“ Subtests 1 und 2 des CFT 1 ( $\Sigma 2$ ) der auf die Verbesserung der Wahrnehmung abzielende Anspruch des DenkMit überprüft werden kann. Wir haben daher diese Subtests ebenfalls vorgegeben, berichten über sie hier jedoch nicht (vgl. HAGER u. HÜBNER 1996), sondern nur über die Ergebnisse in der POD von SAUTER (1979). Unter dem DenkMit von SYDOW und MEINCKE (1994) ist damit

genau wie unter dem Denktraining von KLAUER (1989a) die  $\Sigma 3$  des CFT relevant, und dieser Umstand rechtfertigt es, beide Programme als *Alternativen* (hinsichtlich des induktiven Denkens) zu begreifen und ihre relative Wirksamkeit zu untersuchen (vgl. auch SCHMUDE u. SYDOW 1994). Allerdings: Durch diese AV  $\Sigma 3$  entsteht das bereits angesprochene Problem der „adäquaten“ Erfassung des induktiven Denkens, auf das wir im folgenden noch weiter eingehen.

Um über den Vergleich der beiden Denkförderprogramme (vergleichende Evaluation) hinaus feststellen zu können, ob die beiden Programme unter den von uns realisierten Rahmenbedingungen grundsätzlich wirksam waren (isolierte Evaluation), bildeten wir eine dritte Versuchsgruppe, mit der ebenfalls ein Programm durchgeführt wurde (s.o.). Dieses sollte den Kindern das gleiche Ausmaß an Zuwendung ermöglichen wie die beiden anderen Programme, sie in der gleichen Weise kognitiv beanspruchen und zudem unter den gleichen Randbedingungen durchführbar sein wie die beiden Alternativprogramme, ohne dabei aber irgendwelche Anforderungen in bezug auf das induktive Denken zu stellen oder zu enthalten (Kontrollprogramm). Unsere Wahl fiel auf die *Gedächtnisförderprogramme* der Forschergruppe um MATJUGIN. Aus den Übungsheften zum visuellen Gedächtnis (MATJUGIN u. TSCHAKABERIA 1996) und zum Tastgedächtnis (MATJUGIN et al. 1995) stellten wir Material so zusammen, daß die zuvor genannten Bedingungen so gut wie möglich erfüllt wurden. Dabei wird beim Programm zum Tastgedächtnis mit verschiedenen sogenannten Tastpappen operiert, die wir herstellten und einsetzten, um auch beim Kontrollprogramm die bei den anderen Programmen vorgesehenen Möglichkeiten der motorischen Beschäftigung mit dem Aufgabenmaterial zu realisieren – beim Denktraining werden dazu Bauklötze benutzt und beim DenkMit Puzzleteile.

Obwohl auch dieses Kontrollprogramm ausschließlich mit visuellem Material operiert, enthält es doch keine Aufgaben und keine Aufforderungen zum induktiven Denken im Sinne von Regelableitungen – es zielt auf bestimmte Gedächtnisaspekte ab. Allerdings bedarf es auch bei diesem Programm eines Wahrnehmungsaufwands, der nicht unbedingt als geringer einzuschätzen ist als bei den Denkförderprogrammen – mit dem wohl entscheidenden Unterschied allerdings, daß beim Denktraining das wahrnehmungsmäßige Vergleichen gezielt geübt wird und daß die Wahrnehmungsförderung beim DenkMit eines der Programmziele darstellt. Aus dieser Sicht stellt das Kontrollprogramm, bei dem Prozesse der Wahrnehmung in erster Linie im Hinblick auf die Bewältigung von Gedächtnisaufgaben aktiviert, nicht aber systematisch geübt werden, sozusagen eine „base-line“ im Hinblick auf Wahrnehmungsprozesse dar (vgl. auch HASSELHORN et al. 1995). Soweit wir dies überblicken, ist kaum ein kindgerechtes Programm denkbar, das ohne die Aktivierung von Prozessen der visuellen Wahrnehmung auskommen kann. – Im übrigen setzt auch KLAUER (1995c) in einer neueren Erprobung einer Version des Denktrainings für ältere Kinder ein *Gedächtnistraining* als *Kontrollprogramm* ein.

Zusammenfassend führten wir also einen Vergleich der beiden Alternativprogramme durch, deren grundsätzliche Wirksamkeit in der von uns gewählten Fördersituation jeweils zusätzlich gegen ein anders geartetes Kontrollprogramm festgestellt wurde.



## 2 Weitere abhängige Variablen: kognitionspsychologisch begründete Verfahren

Man kann an den psychometrischen Verfahren CFT und POD grundsätzlich kritisieren, daß sie „ergebnisorientiert“ sind und daß sie damit nicht die Erfassung von *Strategieveränderungen* erlauben. Gerade aber die Übung neuer bzw. die Verbesserung vorhandener Strategien und damit die Beeinflussung von Denkprozessen kann als (weiteres) *zentrales Ziel beider Denkförderprogramme* angesehen werden, und besonders KLAUER (z. B. 1989a, 1991) betont diesen Anspruch für seine Denktrainings wiederholt, konzediert aber (KLAUER 1995a, S. 14), daß er und seine Arbeitsgruppe „wenig getan habe“, um zu überprüfen, inwiefern dieser Anspruch einer empirischen Überprüfung standhält. In dem einen veröffentlichten Versuch von WINDGASSE-FISCHER (in KLAUER 1991, S. 65–66), in dem das Strategieverhalten der Kinder u. a. mittels Protokollen des lauten Denkens untersucht wurde, erwiesen sich die Unterschiede zwischen den verschiedenen trainierten Versuchsgruppen als wenig überzeugend, zumal die Kinder drei (!) Lösungsvorschläge für die ihnen gestellte Aufgaben abgeben konnten (KLAUER 1991, S. 66).

*Konzeptbildungsaufgaben: Zur Erfassung von Strategien* im Bereich des induktiven Denkens wird in der grundlagenorientierten Kognitionspsychologie auf das klassisch zu nennende Paradigma der *Konzeptbildung* zurückgegriffen (vgl. BRUNER et al. 1956; LÜER u. SPADA 1992, S. 236–242). In dieser Tradition ist auch der Wisconsin Card Sorting Test (WCST) (GRANT u. BERG 1948; HEATON 1981) zu sehen. Unter den zahlreichen Untersuchungen mit dem WCST finden sich auch viele mit Kindern, so daß das Material und die Vorgehensweise für Kinder grundsätzlich als geeignet angesehen werden kann. Das Material des WCST besteht aus Karten, auf denen die systematischen Kombinationen dreier Merkmalsdimensionen (Form, Farbe, Anzahl) mit jeweils vier Ausprägungen (Form: Dreieck, Stern, Kreuz, Kreis; Farbe: rot, grün, gelb, blau; Anzahl: eins, zwei, drei, vier) abgebildet sind.

Für unsere Untersuchung wurden die Karten auf die Größe von 7 mal 7 cm verkleinert und in Abweichung vom WCST zur besseren Handhabung auf einen Kartonstreifen geklebt. Analog zum Vorgehen im WCST wurden den Kindern vier Stimulus-Karten (ein rotes Dreieck, zwei grüne Sterne, drei gelbe Kreuze und vier blaue Kreise) vorgelegt. Es wurden nur solche Antwortkarten verwendet, die den Stimuluskarten eindeutig zugeordnet werden können (z. B. teilt die Antwortkarte der Originalversion mit zwei roten Dreiecken mit der ersten Stimuluskarte – ein rotes Dreieck – zwei Merkmale – Farbe und Form – und wurde deshalb aussortiert). Damit reduzierte sich die Anzahl der Antwortkarten auf 24. Die Antwortkarten wurden durchnummeriert und gaben so eine Standardreihenfolge an. Diese wurde derart gewählt, daß zwei aufeinanderfolgende Antwortkarten niemals die gleiche Farbe, Form oder Anzahl aufwiesen. Dadurch wurden alle Antworten eindeutig zuordenbar, und das Verfahren selbst wurde für die Kinder und auch für die Testpersonen einfacher und kürzer. Die Aufgabe für das Kind bestand darin, eine Zuordnungsregel herauszufinden. Das Vorgehen ist u. E. geeignet, das Vorgehen oder die Strategie beim induktiven Erschließen einwertiger Konzepte oder Regeln zu ermitteln.

Als konkrete AVn fungierten dann fünf Maße, die sich aus den KBA und ihrer Durchführung ergaben. Erhoben wurde erstens die „Anzahl der gefundenen Regeln“ (GR),

operationalisiert über sechs richtig zugeordnete Karten in Folge, um die Wahrscheinlichkeit zufällig richtiger Lösungen gering zu halten. Während die Förderung von systematischen Vergleichsprozessen als Lösungsstrategie für induktive Aufgaben ein wesentlicher Anteil in beiden Denkförderprogrammen ist (vgl. KLAUER 1989a, S. 24f.; SYDOW u. MEINCKE 1994, S. 55), wird auf die Begriffe der Regel und der Regelableitung nur im Denktraining für Kinder I (KLAUER 1989a) ausdrücklich eingegangen, und nur dort werden die Kinder angehalten, diese auch zu formulieren. Insofern sind die KBA in bezug auf den Transfer näher am Denktraining als am DenkMit. Daher wird bei der Auswertung der KBA auch zwischen dem Auffinden der Regel, das ein Ergebnis mit beiden Programmen eingeübter induktiver Denkprozesse ist, und ihrer Benennung unterschieden, die nur beim Denktraining thematisiert wird. Die „Anzahl der richtig benannten Regeln“ (BR) fungierte also als zweite AV; nach der Regel wurde jeweils explizit gefragt.

Ferner erhoben wurde die „(relative) Anzahl der richtigen Zuordnungen“ (RZ), in die neben der Anwendung einer gefundenen Regel auch die Stringenz des Vorgehens eingeht. Das Anwenden einer Regel wird ebenfalls besonders beim Denktraining geübt. Als vierte AV wurde die „(relative) Anzahl der sinnlosen Zuordnungen“ (SZ) von Antwortkarten zu den Stimuluskarten betrachtet, also Antwortkarten, die kein Merkmal der Stimuluskarte aufwiesen. Wenn es den Kindern tatsächlich gelingt, sich die in den Programmen vermittelte Strategie des systematischen Vergleichens anzueignen und auf die KBA zu transferieren, dann ist zu erwarten, daß sinnlose Zuordnungen nach der Intervention seltener vorkommen als vorher. Andererseits: Je wahlloser ein Kind vorgeht, desto höher fällt dieser Quotient bei ihm aus. Die AV SZ kann somit als Indikator für planvolles Vorgehen im Sinne einer verbesserten Lösungsstrategie herangezogen werden.

In diesem Sinne ist auch die fünfte AV „Gesamtanzahl der Zuordnungen“ (GZ) zu sehen. Wenn die Kinder (im Vortest) zunächst die Anzahl ihrer Zuordnungen mehr oder minder willkürlich wählen, wobei die individuelle Obergrenze von der Aufmerksamkeits- und Konzentrationsspanne des Kindes abhängen dürfte, dann müßte sich nach einem entsprechenden Strategietraining, das auch oder sogar vor allem die Komponenten reflexives Vorgehen und genaues Arbeiten bzw. Hinschauen enthält, diese Gesamtanzahl innerhalb der individuellen der Aufmerksamkeits- und Konzentrationsspanne des Kindes reduzieren. Auf die Gesamtzahl der Zuordnungen erfolgte jeweils auch die angesprochene Relativierung.

Fraglos sind die AVn der Konzeptbildungsaufgaben hochgradig (entweder positiv oder negativ) miteinander korreliert (vgl. auch Tab. 5 unten); andererseits enthält jede von ihnen Informationen, die in keiner anderen enthalten ist. Aus diesem Grunde erfolgt eine separate Betrachtung aller aufgeführten AVn, die zudem separaten Hypothesen zugeordnet werden (s.u.).

*Analogieaufgaben:* Da die AVn der Konzeptbildungsaufgaben teilweise näher am Denktraining als am DenkMit liegen, versuchten wir, auch einen kognitionspsychologisch begründbaren Aufgabentyp zu verwenden, der näher am DenkMit lag als am Denktraining. Hierfür wählten wir Analogieaufgaben, bei denen jedoch teilweise *negative* Vortest-Nachtest-Korrelationen auftraten, die auf die Ungeeignetheit unseres Materials hinwiesen und über die wir daher nicht weiter berichten.

### 3 Evaluationshypothesen

Bei den zu prüfenden Hypothesen muß zwischen denen unterschieden werden, die sich auf die *isolierte Evaluation* (I) beziehen und damit auf den Vergleiche beider Denkförderprogramme jeweils mit dem Kontrollprogramm (Gedächtnisförderung), und denen, die sich auf die *vergleichende Evaluation* (V) beziehen und damit nur auf die beiden Alternativprogramme. Des weiteren ist zwischen denjenigen Hypothesen zu unterscheiden, die sich aus den Ansprüchen der jeweiligen Programmautoren ergeben *und* die sich auf von ihnen autorisierte AVn ( $\Sigma 3$  im CFT 1) beziehen, und solchen Hypothesen, die sich zwar auf Ansprüche der Programmautoren beziehen, aber über nicht von diesen in Erwägung gezogenen AVn geprüft werden (POD, KBA), und zuletzt einigen konkurrierenden bzw. ergänzenden Hypothesen, die vor allem die mögliche Wahrnehmungsförderung beim Denktraining betreffen. Wir fassen die Vielzahl von einzelnen Hypothesen zu leichter überschaubaren Gruppen zusammen, wobei wir an dieser Stelle nicht über alle möglichen Hypothesen und die sie betreffenden Ergebnisse berichten wollen (vgl. auch HAGER u. HÜBNER 1996).

In der isolierten Evaluation (I) dreht es sich um den Nachweis der speziellen Wirksamkeit der beiden Denkförderprogramme, der für beide Programme getrennt und getrennt für die in Frage kommenden AVn jeweils gegen das Kontrollprogramm über einen Vergleich der Veränderungen vom Vor- zum Nachtest geprüft wird. Als Voraussetzung für die Durchführung dieser Vergleiche müssen die Kontraste innerhalb der drei Interventionsgruppen, also vom Vor- zum Nachtest, selbst vorhergesagetermaßen ausgegangen sein (siehe zu den Einzelheiten HAGER 1995). – Die Hypothesen der *Hypothesengruppe I,1* betreffen den jeweiligen Wirksamkeitsanspruch der Programmautoren und besagen allgemein: Unter den beiden Denktrainings resultiert gegenüber der Kontrollgruppe eine Verbesserung bei der Bearbeitung von induktive Anforderungen stellenden Aufgaben, soweit sie in den vorgesehenen psychometrischen AVn ( $\Sigma 3$ ) erfaßt werden können; beim DenkMit resultiert zusätzlich noch eine Verbesserung auf der Wahrnehmungsseite (POD). Insofern auch mit dem Kontrollprogramm „akzidentelle“ Verbesserungen resultieren (s. o.), sollen diese stets geringer ausfallen als bei den Denkförderprogrammen, die bestimmte Prozesse gezielt einüben wollen, die das Kontrollprogramm nicht gezielt thematisiert. Im einzelnen betrachten wir die folgenden Hypothesen.

*Hypothese I,1-DM*: Das Programm DenkMit (DM), das auf das induktive Denken (ID) und auf die Wahrnehmung (W) abzielt, führt sowohl bei den auf das induktive Denken bezogenen Subtests 3, 4 und 5 ( $\Sigma 3$ ) zu größeren Verbesserungen als das Kontrollprogramm (*Hypothese I,1-DM-ID*( $\Sigma 3$ )) als auch beim Wahrnehmungstest POD (*Hypothese I,1-DM-W*(POD))(Ansprüche der Programmautoren; allerdings wird der Wahrnehmungsbereich über eine von ihnen nicht in Betracht gezogene AV erfaßt).

*Hypothese I,1-DT*: Das Denktraining (DT), das nur auf das induktive Denken (ID) abzielt, führt bei den Subtests 3, 4 und 5 des CFT 1 ( $\Sigma 3$ ) zu größeren Verbesserungen als das Kontrollprogramm (*Hypothese I,1-DT-ID*( $\Sigma 3$ )), nicht aber beim Wahrnehmungstest POD (*Hypothese I,1-DT-NW*(POD))(Ansprüche des Programmautors). – *Hypothese I,1-DT-W*(POD): Obwohl das Denktraining keine Verbesserung auf der Wahrnehmungsseite *anstrebt*, bringt die Fokussierung auf das Üben wahrnehmungsgebundener Vergleichsprozesse es mit sich, daß es auch zu Verbesserungen im Wahrnehmungsbereich führt, erfaßt über die POD (Konkurrenzhypothese zur *Hypothese I,1-DT-NW*(POD); „Wahrnehmungshypothese“ nach HAGER u. HASSELHORN 1993b, 1995b).

Die Hypothesen der *Hypothesengruppe V,1* betreffen die Vergleiche der beiden alternativen Förderprogramme für die beiden psychometrischen AVn. Im Vergleich der beiden alternativen Programme erwarten wir mit KLAUER (1995b) Vorteile seines Denktrainings gegenüber dem DenkMit bei den Subtests 3, 4 und 5 des CFT 1, die von ihrer Struktur her eher Aufgaben im DT als denen im DM entsprechen (*Hypothese V,1-Σ3*). Demgegenüber sollte das DenkMit das Denktraining in der POD übertreffen (*Hypothese V,1-POD,1*). Sofern jedoch mit dem Denktraining auch Wahrnehmungsaspekte merkbar gefördert werden, sollte es auch bei der POD dem DenkMit nicht unterlegen sein (*Hypothese V,1-POD,2*; Konkurrenzhypothese zur *Hypothese V,1-POD,1*).

Die *Hypothesengruppe I,2* bezieht sich auf die Verbesserung der *Strategien* zur Lösung von Aufgaben, die induktive Denkanforderungen stellen; diese Strategien werden mit den *kognitionspsychologischen AVn*, die anhand der Konzeptbildungsaufgaben erhoben wurden, zu erfassen versucht. Generell soll gelten: Mit der Verbesserung der Leistung bei Aufgaben, die induktive Anforderungen stellen, einher geht auch eine Verbesserung der *Strategien* zur Bearbeitung derartiger Aufgaben. Da diese Strategien aber nur mit den Denkförderprogrammen geübt werden, sind im Vortest-Nachtest-Vergleich Verbesserungen nur unter den beiden Denkförderprogrammen zu erwarten, nicht aber beim Kontrollprogramm. Sollten etwa lernbedingt oder „akzidentell“ bei letzterem derartige Verbesserungen auftreten, dann sollen sie durchgängig geringer ausfallen als unter den Denkförderprogrammen. Sollten diese Voraussetzungen eintreffen, werden die Leistungen unter den beiden Denkförderprogrammen getrennt mit denen unter dem Kontrollprogramm verglichen. Es handelt sich hier also erneut um Vergleiche einer *isolierten Evaluation* (I). Im einzelnen ergeben sich daraus folgende Hypothesen für die Vergleiche der Vortest-Nachtest-Kontraste pro Programm (siehe zur Begründung dieser Art der Prüfung HAGER 1995).

*Hypothese I,2-GR(DM)* und *Hypothese I,2-GR(DT)*: Die Anwendung der im DenkMit (DM) bzw. im Denktraining (DT) vermittelten induktiven Denkstrategien führt zu einer Erhöhung der Anzahl der gefundenen Regeln (GR). – *Hypothese I,2-BR(DM)*: Da im DenkMit das Benennen von Regeln nicht geübt wird, unterscheidet es sich in seiner Wirkung auf die Anzahl richtig benannter Regeln (BR) nicht vom Kontrolltraining. – *Hypothese I,2-BR(DT)*: Das Denktraining fördert auch das Benennen von Regeln und führt daher zu einer größeren Anzahl der richtig benannten Regeln (BR) als das Kontrolltraining. – *Hypothese I,2-RZ(DM)* und *Hypothese I,2-RZ(DT)*: Durch die Förderung induktiver Prozesse gelingt die Anwendung von Regeln besser, und somit führt das DenkMit bzw. das DT zu einer größeren (relativen) Anzahl der richtigen Zuordnungen (RZ) als das Kontrolltraining. – *Hypothese I,2-SZ(DM)* und *Hypothese I,2-SZ(DT)*: Das DenkMit bzw. das Denktraining führen durch entsprechende Übungselemente zu einer Verbesserung der Strategien beim Bearbeiten induktiver Aufgaben. Daher nimmt unter beiden die (relative) Anzahl der „sinnlosen“ Zuordnungen (SZ) im Vergleich zum Kontrolltraining ab. – *Hypothese I,2-GZ(DM)* und *Hypothese I,2-GZ(DT)*: Die Anwendung der im DenkMit bzw. im Denktraining vermittelten Denkstrategien führt zu einer Verminderung der Gesamtanzahl der Zuordnungen (GZ) im Vergleich zum Kontrollprogramm, in dem keine Änderung auftreten soll.

Die *Hypothesengruppe V,2* bezieht sich auf den Vergleich der Verbesserungen der *Strategien* zur Lösung von Aufgaben, die induktive Denkanforderungen stellen bei den beiden Denkförderprogrammen (*vergleichende Evaluation*; V). Generell ist dabei zu erwarten: Da das Denktraining einen weiteren Bereich von Strategien einübt als das DenkMit und da das Auffinden von Regeln bei ihm immer wieder betont wird, soll es zu größeren Strategieverbesserungen führen als das DenkMit. Im einzelnen heißt dies für die möglichen Vortest-Nachtest-Vergleiche:

*Hypothese V,2-GR*: Da im Denktraining beim Einüben induktiver Denkstrategien auf Regelableitungen Bezug genommen wird, werden hier mehr Regeln gefunden (GR) als nach der Durchführung des DenkMit. – *Hypothese V,2-BR*: Im Denktraining wird die Formulierung von Regeln geübt, nicht aber im DenkMit. Daher ist für das Denktraining eine größere Anzahl richtig benannter Regeln (BR) zu erwarten als im DenkMit. – *Hypothese V,2-RZ*: Beide Programme fördern induktive Denkprozesse, aber die zu erbringende Transferleistung zwischen Trainings- und Testaufgaben ist für das Denktraining geringer, weshalb die Anzahl der richtigen Zuordnungen (RZ) nach dem Denktraining größer sein sollte als nach dem DenkMit. – *Hypothese V,2-SZ*: In beiden Programmen wird das systematische Vergleichen als eine für die Bearbeitung induktiver Anforderungen stellender Aufgaben relevanten Lösungsstrategie geübt. Die Anzahl der „sinnlosen“ Zuordnungen sollte in beiden Trainingsgruppen nach der Intervention nicht voneinander abweichen. – *Hypothese V,2-GZ*: In beiden Denkprogrammen werden zwar Lösungsstrategien geübt, wodurch eine Verminderung der Gesamtzahl von Zuordnungen unter beiden zu erwarten ist; es gibt aber keinen Hinweis, der zu der Vermutung Anlaß geben könnte, daß dieses Absinken unter einem Programm stärker ausfällt als unter dem anderen. Dementsprechend erwarten wir bei der AV GZ keine Unterschiede zwischen den Programmen.

Die den jeweiligen Hypothesen zugeordneten und auf den Versuchsplan bezogenen Vorhersagen sowie die ihnen entsprechenden statistischen Hypothesen (Null- oder Alternativhypothesen über Mittelwerte  $\mu$ ;  $H_0$  und  $H_1$ ) sind den Ergebnistabellen unten zu entnehmen.

#### 4 Versuchsdurchführung, Trainingspersonen und trainierte Kinder

Durchgeführt wurden die Programme von Studenten der Psychologie im zweiten Studienabschnitt. Die Programme wurden mit den Trainingspersonen in einer Lehrveranstaltung besprochen. Die freiwillig an der Intervention Teilnehmenden konnten sich dann entscheiden, ob sie das Denktraining oder aber das DenkMit und jeweils das Kontrollprogramm durchführen wollten. Gemäß der sich ergebenden Aufteilung wurden separate Gruppen gebildet, die eine detaillierte Unterweisung in die Programme erhielten und die die Programme auch mehrfach in Ausschnitten miteinander durchführten. Dabei erhielten die Personen, die sich für das DenkMit entschieden hatten, eine umfangreich überarbeitete Version des von den Programmautoren vorgesehenen Manuals. Die Erstellung eines solchen Manuals hatte sich in einem früheren Versuch (HASSELHORN u. HAGER 1995) als erforderlich erwiesen, da viele der Ausführungen der Programmautoren entweder zu Mißverständnissen Anlaß gaben oder aber inkonsistent schienen. Da das Material zum DenkMit noch nicht im Handel erhältlich ist, arbeiteten wir mit Farbkopien des

Originalmaterials, das uns dankenswerterweise Herr Prof. Dr. H. Sydow zur Verfügung gestellt hat.

Obwohl auch PERLETH (1991, S. 270) moniert, daß die konkreten Durchführungsanleitungen zum Denktraining „einen Schwachpunkt darstellen“, haben wir für dieses Programm kein separates Manual erstellt, weil dies vom Programmautor nicht vorgesehen ist. Die Trainingspersonen für dieses Programm erhielten ihre Unterweisung anhand des Manuals des Programmautors (KLAUER 1989a), wobei für etliche der Übungssitzungen mit den Trainingspersonen gezielt solche Aufgaben ausgesucht wurden, deren Bearbeitung sich in früheren Erprobungen des Denktrainings als problematisch erwiesen hatte (HAGER u. HASSELHORN 1993a, 1993b). Daneben erhielten die Trainingspersonen die betreffenden Seiten aus dem Manual zum Durcharbeiten sowie einen vollständigen Satz von Schwarzweißkopien des Programm-Materials, damit sie sich jeweils genau auf die einzelnen Trainingssitzungen vorbereiten konnten. Allen Trainings- und Testpersonen wurde darüber hinaus kurzzeitig auch das Originalmaterial zur Verfügung gestellt. Die vorgesehene Vermittlungsmethode war das gelenkte Entdeckenlassen, wobei bedarfsweise (bei schwierigeren Aufgaben und/oder bei schwächer befähigten Kindern) auch die Methode der verbalen Selbstinstruktion eingesetzt wurde (vgl. auch KLAUER 1989a, S. 97-100).

Die Unterweisung aller Trainingspersonen in das Kontrollprogramm erwies sich als wesentlich einfacher durchführbar, da jede Aufgabe mit einer konkreten Instruktion verbunden ist. Auch besprochen wurde die Durchführung der Vor- und der Nachtest-Aufgaben (s.o.), wobei die Trainingspersonen in die Hintergründe und die Durchführung der einzelnen Aufgaben (psychometrische Tests und Denkaufgaben) eingeführt wurden. Es wurde darauf geachtet, daß keine der insgesamt 18 Trainingspersonen ein Kind testete, das sie später auch trainieren sollte bzw. vorher trainiert hatte. Während der Intervention fand eine Supervision der Trainierenden statt, um sicherzustellen, daß die Programme gemäß den Vorgaben ihrer Autoren durchgeführt wurden.

Für die Teilnahme an der Intervention erklärten sich die Eltern von 52 Kindern bereit, von denen die überwiegende Mehrheit aus den verschiedensten Gründen für ein Jahr von der regulären Einschulung zurückgestellt worden war und den Schulkindergarten besuchte; den Gründen für diese Zurückstellungen (allgemein: Leistungsdefizite, Behinderungen usw.) gingen wir nicht im einzelnen nach, da ja keine gezielte, sondern eine allgemeine Förderung erfolgen sollte. Einige wenige Kinder befanden sich bereits in ersten Klassen, wurden aber von der jeweiligen Klassenleiterin als besonders schwache und daher förderungsbedürftige Kinder ausgewählt. Von den 52 Kindern erwiesen sich zwei wegen sprachlicher Probleme mit den vorgesehenen Materialien als nicht trainierbar; für sie stellten die Trainingspersonen nach Absprache mit den Schulkindergartenleiterinnen und uns individuelle Förderaufgaben zusammen; dies galt auch für zwei weitere wegen geistiger Behinderung bereits vorher aus der regulären Förderung herausgenommene Kinder. Insgesamt verblieben also zunächst  $N=50$  Kinder, die zufällig auf die drei Versuchsgruppen aufgeteilt wurden. Ein weiteres Kind fiel allerdings kurz nach Beginn der Intervention durch einen längeren Krankenhausaufenthalt aus der Förderung, so daß schließlich  $N=49$  Kinder im Altersbereich von 5;10 bis 8;03 Jahren verblieben. Für diese ergaben sich vor der Intervention folgende Altersverteilungen innerhalb der drei Versuchsgruppen: Denktrainingsgruppe ( $n=16$ ): 5;10 bis 7;10 Jahre (Mittelwert: 6;09 Jahre); DenkMit-Gruppe ( $n=17$ ): 6;00 bis 8;03 Jahre (Mittelwert: 6;11 Jahre), und Gedächtnistrainings-Gruppe ( $n=16$ ): 6;01 bis 7;10 Jahre (Mittelwert: 6;08 Jahre). Daß einige der Kinder älter waren als der für das DenkMit vorgesehene Altersbereich (fünf bis sechs Jahre), halten wir für nicht problematisch, da es sich durchgängig um schwächer befähigte Kinder handelte, so daß ihr *Entwicklungsalter* dem vorgesehenen Alter dieses Programms ungefähr entsprechen haben dürfte. Darüber hinaus traten auch bei diesen Kindern keinerlei Deckeneffekte oder sonstige Besonderheiten auf, die auf eine mangelnde Geeignetheit der Programme für sie hindeuteten.

## 5 Ergebnisse und Hypothesenprüfungen

Als Ausgangspunkt zur Prüfung der Hypothesen für die beiden psychometrischen AVn geben wir zunächst in einer Übersichtstabelle die Mittelwerte (M) und Standardabweichungen (s) sowie die Vortest-Nachtest-Korrelationen ( $r_{VN}$ ) für die drei Versuchsgruppen an (Tab. 1:  $\Sigma 3$  des CFT 1, Gesamt-IQ und POD). Zum Zeitpunkt des Nachtests unmittelbar vor Weihnachten 1995 standen leider wegen Krankheit nicht alle Kinder bzw. nicht alle Kinder für alle Erhebungen zur Verfügung, so daß sich unterschiedlich große Stichproben ergaben.

Tab. 1: Ergebnisse in den Subtests 3, 4 und 5 ( $\Sigma 3$ ) und IQ-Werte im CFT 1 von WEISS und OSTERLAND (1980) und in der POD von SAUTER (1979) (Mittelwerte M, Streuungen s und Vortest-Nachtest-Korrelationen  $r_{VN}$ )

	Denktraining (n = 16)		DenkMit (n = 17)		Gedächtnispro- gramm (n = 16)	
	M	s	M	s	M	s
a) $\Sigma 3$ des CFT 1 (Subtests 3, 4 und 5)						
V:	12.00	4.73	14.18	6.22	14.13	5.34
N:	17.69	6.36	19.76 <sup>a</sup>	6.10	20.06 <sup>b</sup>	5.52
$r_{VN}$ :	.70		.54		.84	
b) Gesamt-IQ im CFT 1						
V:	86.19	9.88	91.25	14.64	91.81	10.35
N:	94.00	14.72	98.38 <sup>a</sup>	15.45	101.19 <sup>b</sup>	15.34
$r_{VN}$ :	.66		.78		.78	
c) POD						
V:	15.25	6.26	18.00	7.52	17.67	6.02
N:	20.50	4.73	19.82	6.65	20.07 <sup>c</sup>	5.16
$r_{VN}$ :	.69		.91		.66	

*Anmerkungen:*

<sup>a</sup> Die Nachtestwerte beruhen auf n = 16.

<sup>b</sup> Den Nachtestwerten liegen n = 14 zugrunde.

<sup>c</sup> Die Nachtestwerte beruhen auf n = 15.

Als Voraussetzungen für die anschließenden Tests wurde dann das Nicht-Abfallen bzw. das Ansteigen der Leistungen vom Vor- zum Nachtest anhand von t-Tests für abhängige Stichproben geprüft; die dabei benutzte Streuung ist in der Spalte „ $s_{\text{Test}}$ “ angegeben (vgl. Tab. 2). In allen Fällen erwiesen sich die Ergebnisse als vorhersagekonform und teilweise auch mit recht großen Effekten  $d_{\text{emp}}$  verbunden; die inhaltlichen Voraussetzungen für die Interpretierbarkeit der nachfolgenden Tests in Tabelle 3 ist damit gegeben (vgl. dazu HAGER 1995).

Tab. 2: Ergebnisse der Tests zur Prüfung der Leistungsveränderungen vom Vor- zum Nachtest im CFT und in der POD sowie empirische Effektgrößen

Programme	Vorhersage	$t_{emp}$	$s_{Test}$	df	sign.?	$d_{emp}$	vorhersagekonform?
a) CFT 1 ( $\Sigma 3$ , Subtests 3, 4 und 5)							
DT	Anstieg (H1)	4.99	4.56	15	ja	1.0	ja
DM	Anstieg (H1)	3.89	5.92	16	ja	0.9	ja
KP	Nicht-Abfallen (H0)	7.75	3.07	15	nein	(1.1)	ja
b) POD							
DT	Nicht-Abfallen (H0)	4.60	4.57	15	nein	(0.9)	ja
DM	Anstieg (H1)	2.45	3.07	16	ja	0.3	ja
KP	Nicht-Abfallen (H0)	2.00	4.66	14	nein	(0.4)	ja

*Anmerkungen:*

DT: Denktraining; DM: DenkMit; KP: Kontrollprogramm;  $\alpha = .05$  für alle Tests. – In Klammern ist hinter jedem Programm für die jeweilige AV angegeben, ob ein Anstieg der Testleistungen vom Vor- zum Nachtest erwartet wird oder lediglich ein Nicht-Abfallen. Ebenfalls angegeben ist der Typ der statistischen Hypothese ( $H_0$  oder  $H_1$ ; vgl. Text), die der jeweiligen Vorhersage entspricht. –  $s_{Test}$  bezeichnet die zum Testen herangezogene Streuung der Mittelwertsdifferenz bei abhängigen Daten. –  $d_{emp}$ : auf die mittlere Binnenstreuung standardisierte Mittelwertsdifferenz. – Bei beibehaltenen Nullhypothesen erübrigt sich die Berechnung von Effektgrößen in der von der Alternativhypothese definierten Richtung, da sie nur zufällige Abweichungen von Null anzeigen. Allerdings können mit beibehaltenen vorhersagekonformen und *gerichteten* Nullhypothesen beträchtliche Effekte in der *vorhersagekonformen* Richtung verbunden sein, die hier der Information halber (eingeklammert) angegeben werden.

Zur Prüfung der Evaluationshypothesen wurden dann geplante t-Tests über Differenzen von Mittelwertsdifferenzen (Interaktionen Programm mal Erhebungszeitpunkte) auf der Basis von varianzanalytisch bestimmten Residualvarianzen durchgeführt; deren Wurzel ist jeweils wieder in der Spalte „ $s_{Test}$ “ angegeben. Das heißt, es wurde geprüft, ob sich die Veränderungen vom Vor- zum Nachtest innerhalb der jeweiligen Versuchsgruppen zwischen ihnen voneinander unterschieden. Die Fehlerwahrscheinlichkeit  $\alpha$  wurde durchgängig auf  $\alpha = .05$  festgesetzt. Eine Adjustierung erfolgte nicht, da jede psychologische Hypothese letztlich mit nur *einem* Test überprüft wurde (vgl. dazu ausführlich HAGER 1995). Wir verzichten aus Raumgründen auf die Darstellung der *Testplanung* zur Kontrolle der beiden statistischen Fehlerwahrscheinlichkeiten  $\alpha$  und  $\beta$  und zum Nachweis von statistischen Effekten bestimmter Größe. Die Ergebnisse der Vergleiche zwischen den Programmen sind in Tabelle 3 aufgeführt.

Die Hypothesen der isolierten Evaluation der Gruppe I,1 haben sich *nicht* bewähren können, da die Leistungszuwächse beim Kontrollprogramm *erwartungswidrig* denen unter den Denkförderprogrammen nicht nachstanden: Unter dem Kontrollprogramm sind bei den induktive Anforderungen stellenden Subtests des CFT 1 ( $\Sigma 3$ ) Verbesserungen in der gleichen Größenordnung aufgetreten wie unter den Denkförderprogrammen. Hinsichtlich der Wahrnehmungsaspekte, erfaßt über die POD, haben sich unter dem Denktraining größere Verbesserungen ergeben als unter dem Kontrollprogramm; dies spricht für eine weitere Bewährung der „Wahrnehmungshypothese“ von



Tab. 3: Ergebnisse der Tests zur Prüfung der Evaluationshypothesen  
(Vergleich der Vortest-Nachtest-Kontraste)

Hypothese und AV	Vorhersage	$t_{emp}$	$s_{Test}$	df	sign.?	$d_{emp}$	Hypothese bestätigt?
I,1-DM-ID( $\Sigma 3$ ):	DM > KP	0.20	3.36	31	nein	----	nein
I,1-DT-ID( $\Sigma 3$ ):	DT > KP	0.17	2.74	30	nein	----	nein
I,1-DM-W(POD)	DM > KP	-0.41	2.75	30	nein	----	nein
I,1-DT-NW(POD):	DT=KP	1.72	3.26	29	ja	(0.5)	nein
I,1-DT-W(POD):	DT > KP	1.72	3.26	29	ja	0.5	ja
V,1- $\Sigma 3$ :	DT > DM	0.0	3.74	31	nein	----	nein
V,1-POD,1:	DM > DT	-2.54	2.73	31	nein	(-0.5)	nein
V,1-POD,2:	DM $\leq$ DT	-2.54	2.73	31	ja	-0.5	ja

*Anmerkungen:*

DT: Denktraining, DM: DenkMit, KP: Kontrollprogramm. – Die A-priori-Vergleiche der Kontraste wurden per Interaktionstest (einseitiger t-Test) einer Meßwiederholungsvarianzanalyse durchgeführt; die jeweils zum Testen benutzte Residualstreuung ( $s_{Test}$ ) und ihre Freiheitsgrade (df) sind angegeben;  $\alpha=.05$  für alle Tests. – Zu weiteren Erläuterungen vgl. die vorhergehende Tabelle.

HAGER und HASSELHORN (1993b, 1995) selbst im Vortest-Nachtest-Vergleich, obwohl sie eigentlich in erster Linie als „Kompetenzhypothese“ verstanden wird, die man aber erst unter Einbezug der Follow-up-Untersuchung prüfen kann (vgl. zu Einzelheiten HASSELHORN u. HAGER 1996). Die Bewährung der Wahrnehmungshypothese widerspricht gleichzeitig besonders KLAUERS diesbezüglicher Auffassung, da das Denktraining auch im Vergleich zum DenkMit, das ja gezielt die Wahrnehmung fördern will, zu größeren Verbesserungen führte. Dieser Befund und der Umstand, daß beim DenkMit keine größeren Verbesserungen resultierten als unter dem Kontrollprogramm, spricht dafür, daß der diesbezügliche Anspruch der Autoren des DenkMit nicht als erfüllt angesehen werden kann: Die mit diesem Programm erreichbaren Verbesserungen bewegen sich in der Größenordnung von denen, die unter einem Programm aufgetreten sind, mit dem Wahrnehmungsleistungen nicht gezielt angesprochen werden, sondern bestenfalls „akzidentell“ im Hinblick auf die Verbesserung von Gedächtnisleistungen. Bei der vergleichenden Evaluation hat sich darüber hinaus die Hypothese KLAUERS, die eine Überlegenheit des Denktrainings gegenüber dem DenkMit behauptet (Hypothese V,1- $\Sigma 3$ ),

nicht bewähren können: Die Leistungssteigerungen in den beiden Alternativprogrammen hinsichtlich der induktive Anforderungen stellenden Aufgaben der  $\Sigma 3$  unterscheiden sich nicht. Entscheidend ist jedoch, daß sich diese Leistungen auch nicht von denen unter dem Kontrollprogramm unterscheiden (isolierte Evaluation). Dies bedeutet, daß sich in unserer Untersuchung *die beiden Denkförderprogramme als gleich unwirksam erwiesen haben*, was die psychometrische AV  $\Sigma 3$  im CFT 1 betrifft.

Eine naheliegende Erklärung für dieses Resultat könnte darin bestehen, daß das Kontrollprogramm von den beiden Denkförderprogrammen nicht hinreichend verschieden gewesen ist (vgl. dazu KLAUER 1995b). Dagegen spricht jedoch, daß als Kontrollprogramm nur solche Aufgaben überwiegend zum Tastgedächtnis zusammengestellt worden waren, die keine induktiven Anforderungen an die Kinder stellten und bei denen dies auch durch die Instruktion in keiner Weise nahegelegt wurde (s.o.). Demgegenüber zielen die beiden Denkförderprogramme gerade auf dieses Denken ab. Auch nahe liegt der Hinweis auf die mangelhafte Implementation der Programme, doch steht diesem Argument die Tatsache der gründlichen Vorbereitung der Trainingspersonen und die während der Intervention stattfindende Supervision entgegen (s.o.). Dies erschwert es, den „Mißerfolg“ den Trainingspersonen anzulasten (vgl. zu diesem Argument auch KLAUER 1995c).

Es scheint daher ein Faktor wirksam geworden zu sein, der offensichtlich nicht mit den spezifischen Zielen der Autoren identisch ist, sondern der unter allen drei Programmen sehr wirksam war. Die übliche Bemerkung in diesem Zusammenhang bezieht sich vor allem auf die Entwicklung und/oder Reifung und auf Wirkungen der Retestung (z. B. KLAUER 1993b). Diese Wirkungen sind aber besonders bei randomisierter Zuweisung der Kinder zu den Programmen als für alle Versuchsgruppen gleich anzunehmen, wirken also nicht differentiell und damit nicht validitätsmindernd. Zudem ist bei der Kürze der Zeit vom Vor- zum Nachtest nicht mit großen Wirkungen der Entwicklung und/oder Reifung zu rechnen, wie die Inspektion der Normentabellen in WEISS und OSTERLAND (1980) belegt. Abgesehen davon, zeigte sich auch für die IQ-Werte in allen drei Versuchsgruppen ein signifikanter Anstieg vom Vor- zum Nachtest (vgl. auch Tab. 2). Nun verbleiben noch (mindestens) zwei weitere Faktoren, die zu der Leistungsverbesserung in den Gruppen beigetragen haben können. Zum einen handelt es sich um die *kognitive Beanspruchung* unter den drei Interventionen. Wie oben erwähnt, ist in unserem Fall für das Kontrollprogramm unter anderem zu fordern, daß es den beiden interessierenden Programmen vergleichbar kognitiv beanspruchend ist, dabei aber auf einen anderen Bereich abzielt als die zu evaluierenden Programme. Die entsprechende Befragung der Trainerinnen und Trainer, die jeweils eines der Denkförderprogramme und das Kontrollprogramm durchführten, ergab jedoch keinen Hinweis auf differentielle kognitive Belastungen der Kinder.

Zum anderen könnte noch ein weiterer Faktor zu unseren Resultaten geführt haben: nämlich die mit einer Einzelintervention unvermeidbar verbundene *Zuwendung* oder auch soziale Interaktion zwischen den Trainingspersonen und den Kindern. Die Wirkung dieser Zuwendung kann nämlich nach OERTER (1973) darin bestehen, daß die Kinder sich allgemeine Bearbeitungsstrategien besonders im Sinne eines allgemein reflexiveren Vorgehens aneignen und daß sie besonders motiviert werden. Die Zuwendung kann also zu einer unspezifischen kognitiven Förderung bei bestimmten „strategieanfälligen“ AVn führen, die ihrerseits Leistungssteigerungen nach sich zieht. Vergleicht man unsere Ergebnisse aus diesem Blickwinkel mit denen von HAGER und HASSELHORN (1995), die beim Denktraining ebenfalls beträchtliche Zuwendungswir-

kungen nachweisen konnten, dann drängt sich als der plausibelste Kandidat für eine Erklärung unserer Befunde bei der psychometrischen Erfassung des induktiven Denkens die *interventionsspezifische Zuwendung* auf. Es sei an dieser Stelle auch hinzugefügt, daß die teilweise großen Effekte, die KLAUER (1989a, 1991) für sein Denktraining nachgewiesen hat, immer auf einen Vergleich mit einer *nicht-trainierten* Versuchsgruppe zurückgehen, in der natürlich keine besondere Zuwendung auftreten kann.

Der interventionsspezifischen Zuwendung scheint also bei Betrachtung von psychometrischen AVn eine größere Bedeutung zuzukommen, als es die Programmautoren zu konzedieren bereit sind. Aber dennoch haben die Denkförderprogramme etwas bewirkt, das über Zuwendungswirkungen hinausgeht. Dies wird deutlich, wenn man die Konzeptbildungsaufgaben betrachtet, denen wir uns nun zuwenden.

Die Tabelle 4 enthält die Ergebnisse zu den AVn der Konzeptbildungsaufgaben (KBA), Tabelle 5 die Interkorrelationen zwischen ihnen zu den beiden Erhebungszeitpunkten, Tabelle 6 die Vortest-Nachtest-Vergleiche und Tabelle 7 die Vergleiche der Veränderungen über die Programme.

Tab. 4: Ergebnisse in den Konzeptbildungsaufgaben (KBA; Mittelwerte M, Streuungen s und Vortest-Nachtest-Korrelationen  $r_{VN}$ )

AVn	Denktraining (n=14)		DenkMit (n=16)		Gedächtnisprogramm (n=15)		
	M	s	M	s	M	s	
GR	V:	2.71	1.44	2.56	1.63	2.53	1.64
	N:	3.79	2.58	4.38	1.96	4.60	1.72
	$r_{VN}$ :	.58		.58		.23	
BR	V:	1.71	1.54	1.38	1.54	1.07	1.67
	N:	3.00	2.75	3.38	2.58	2.60	2.50
	$r_{VN}$ :	.75		.75		.49	
RZ	V:	.47	.17	.40	.17	.45	.21
	N:	.60	.26	.59	.25	.55	.16
	$r_{VN}$ :	.51		.57		.21	
SZ	V:	.15	.09	.14	.09	.13	.09
	N:	.11	.08	.09	.09	.14	.09
	$r_{VN}$ :	.61		.13		.02	
GZ	V:	74.71	27.65	85.75	20.13	76.13	28.72
	N:	67.07	29.06	70.06	24.82	76.27	20.07
	$r_{VN}$ :	.26		.38		-.10	

*Anmerkungen:*

GR: Anzahl gefundener Regeln; BR: Anzahl richtig benannter Regeln; RZ: relative Anzahl richtiger Zuordnungen; SZ: relative Anzahl sinnloser Zuordnungen; jeweils bezogen auf die Gesamtzahl der Zuordnungen (GZ).

Tab. 5: Interkorrelationen  $r$  zwischen den AVn der Konzeptbildungsaufgaben zu den beiden Erhebungszeitpunkten (obere Hälfte der Matrix: Vortest; untere Hälfte: Nachtest)

AV	GR	BR	RZ	SZ	GZ
GR		.668	.605	-.400	-.180
BR	.786		.420	-.259	-.100
RZ	.705	.750		-.526	-.760
SZ	-.512	-.538	-.788		.476
GZ	-.230	-.440	-.760	.675	

*Anmerkungen:*

GR: Anzahl der gefundenen Regeln; BR: Anzahl der richtig benannten Regeln; RZ: relative Anzahl der richtigen Zuordnungen; SZ: relative Anzahl der sinnlosen Zuordnungen; GZ: Gesamtzahl aller Zuordnungen.

Tab. 6: Ergebnisse der Tests zur Prüfung der Leistungsveränderungen vom Vor- zum Nachtest bei den fünf AVn der Konzeptbildungsaufgaben (KBA)

AV	Programm	Vorhersage	$t_{emp}$	$s_{Test}$	df	sign.?	$d_{emp}$	vorher-konform?
GR	DT	Ansteigen ( $H_1$ )	1.92	2.09	13	ja	0.5	ja
	DM	Ansteigen ( $H_1$ )	4.31	1.68	15	ja	1.0	ja
	KP	Nicht-Abfallen ( $H_0$ )	3.84	2.09	14	nein	1.2	ja
BR	DT	Ansteigen ( $H_1$ )	2.53	1.90	13	ja	0.6	ja
	DM	Nicht-Abfallen ( $H_0$ )	4.57	1.75	15	nein	(0.9)	ja
	KP	Nicht-Abfallen ( $H_0$ )	2.66	2.23	14	nein	(0.7)	ja
RZ	DT	Ansteigen ( $H_1$ )	2.11	0.23	13	ja	0.6	ja
	DM	Ansteigen ( $H_1$ )	3.56	0.21	15	ja	0.9	ja
	KP	Nicht-Abfallen ( $H_0$ )	1.70	0.24	14	nein	----	ja
SZ	DT	Abfallen ( $H_1$ )	-1.81	0.08	13	ja	-0.4	ja
	DM <sup>a</sup>	Abfallen ( $H_1$ )	----	0.12	15	----	----	(nicht geprüft)
	KP	Nicht-Ansteigen ( $H_0$ )	0.39	0.12	14	nein	----	ja
GZ	DT <sup>a</sup>	Abfallen ( $H_1$ )	----	34.57	13	nein	----	(nicht geprüft)
	DM	Abfallen ( $H_1$ )	-2.48	25.26	15	ja	-0.7	ja
	KP	Nicht-Ansteigen ( $H_0$ )	-0.01	36.62	14	nein	----	ja

*Anmerkungen:*

GR: Anzahl gefundener Regeln; BR: Anzahl richtig benannter Regeln; RZ: relative Anzahl richtiger Zuordnungen; SZ: relative Anzahl sinnloser Zuordnungen; GZ: Gesamtzahl der Zuordnungen;  $\alpha = .05$ . – <sup>a</sup> Wegen der geringen Vortest-Nachtest-Korrelationen (vgl. Tab. 4) werden diese Veränderungen statistisch nicht geprüft. – Zu weiteren Erläuterungen siehe Tabelle 2.

Tab. 7: Vergleich der Vortest-Nachtest-Kontraste der isolierten Evaluation für die fünf AVn der KBA

Hypothese	Vorhersage	$t_{emp}$	$s_{Test}$	df	sign.?	Hypothese bestätigt?
1,2-GR(DM)	DM > KP ( $H_1$ )	0.37	1.33	29	nein	nein
1,2-BR(DM)	DM=KP ( $H_0$ )	-0.65	1.41	29	nein	ja
1,2-RZ(DM)	DM > KP ( $H_1$ )	-1.02	0.14	29	nein	nein
1,2-SZ(DM) <sup>a</sup>	DM > KP ( $H_1$ )	----	0.10	29	----	(nicht geprüft)
1,2-GZ(DM)	DM > KP ( $H_1$ )	1.41	22.11	29	nein	nein
1,2-GR(DT)	DT > KP ( $H_1$ )	-1.28	1.48	27	nein	nein
1,2-BR(DT)	DT > KP ( $H_1$ )	0.32	1.47	27	nein	nein
1,2-RZ(DT)	DT > KP ( $H_1$ )	-0.28	0.17	27	nein	nein
1,2-SZ(DT)	DT > KP ( $H_1$ )	-1.29	0.10	27	nein	nein
1,2-GZ(DT) <sup>a</sup>	DT > KP ( $H_1$ )	----	25.20	27	----	(nicht geprüft)

*Anmerkungen:*

GR: Anzahl gefundener Regeln; BR: Anzahl richtig benannter Regeln; RZ: relative Anzahl richtiger Zuordnungen; SZ: relative Anzahl sinnloser Zuordnungen; GZ: Gesamtanzahl der Zuordnungen. –  $s_{Test}$ : für den t-Test benutzte Teststreuung (Wurzel aus der varianzanalytischen Residualvarianz);  $\alpha = .05$ . – <sup>a</sup> Die betreffenden Hypothesen werden nicht geprüft, weil entweder die Voraussetzungen für die inhaltlich sinnvolle Interpretation dieser Vergleiche, nämlich vorhersagekonforme Vortest-Nachtest-Veränderungen, nicht eingetreten sind, oder weil die Vortest-Nachtest-Korrelationen zu gering ausfielen (vgl. Text). – Wegen der Beibehaltung der vorhersagekonträren Nullhypothesen (und der kleinen t-Werte) wurden keine Effektgrößen  $d_{emp}$  berechnet (vgl. zur Begründung HAGER, 1987). – Zu weiteren Erläuterungen siehe Tabelle 2.

Die Interkorrelationen zwischen den AVn der KBA enthalten keine auffälligen Werte, sondern bewegen sich insbesondere hinsichtlich der Vorzeichen im erwarteten Bereich. – Bezüglich der Vortest-Nachtest-Vergleiche ergeben sich erwartungswidrige Resultate nur bei der AV SZ für das DenkMit, unter dem auch die Vortest-Nachtest-Korrelation sehr gering ausfällt ( $r_{VN} = .13$ ) und die AV GZ für das Denktraining ( $r_{VN} = .26$ ); immerhin weisen aber auch diese Ergebnisse in die vorhergesagte Richtung. Aber da der vorhersagekonforme Ausgang der Testung der Vortest-Nachtest-Kontraste innerhalb der Programme als Voraussetzung der Prüfung der Evaluationshypothesen angesehen wird (vgl. HAGER 1995) und da die Korrelationen vergleichsweise gering (wenn auch positiv) ausfallen, werden die Hypothesen 1,2-SZ(DM) und 1,2-GZ(DT) sowie V,2-SZ und V,2-GZ *nicht* geprüft. Die übrigen Hypothesen der Gruppe 1,2 einer isolierten Evaluation können sich – mit Ausnahme der Hypothese 1,2-BR(DM) – nicht bewähren (vgl. Tab. 7). Erwähnenswert ist aber die Nullkorrelation unter dem Kontrollprogramm (vgl. Tab. 4), die ja andeutet, daß die Variable GZ unter diesem Programm keine systematische Beeinflussung erfahren hat; auch beim DenkMit fallen die Korrelationen  $r_{VN}$  für die AVn SZ und GZ sowie beim Denktraining nur für die AV GZ gering aus.

Im wesentlichen decken sich diese Befunde also mit den Ergebnissen im CFT 1, für den allerdings Anstiege unter *allen* Programmen zu verzeichnen waren; insofern stellt sich das Befundmuster bei den KBA insgesamt doch etwas anders dar. Bevor wir auf

diese Unterschiede eingehen, sollen noch kurz die Hypothesen der Gruppe V,2 (vergleichende Evaluation) betrachtet werden.

In der vergleichenden Evaluation der Denkförderprogramme ist zu erwarten, daß das Denktraining dem DenkMit sowohl beim Auffinden von Regeln (Hypothese V,2-GR) als auch bei ihrer Formulierung (Hypothese V,2-BR) überlegen ist, da beide Aspekte mit dem Programm explizit gefördert werden. Auch die Anwendung der gefundenen Regeln sollte nach dem Denktraining besser gelingen als nach dem DenkMit, da für die Kinder des letzteren eine größere Transferleistung nötig ist, da sie die Anwendung nicht gezielt üben (Hypothese V,2-RZ). Allerdings erwies sich das Denktraining erwartungswidrig dem DenkMit als nur tendenziell, nicht aber statistisch signifikant überlegen (vgl. Tab. 8), so daß sich diese Hypothesen ebenfalls nicht bewähren können, wenn man als vorrangiges Kriterium die statistische Signifikanz bewertet.

Tab. 8: Vergleich der Vortest-Nachtest-Kontraste der vergleichenden Evaluation für die AVn der KBA

Hypothese	Vorhersage	$t_{emp}$	$s_{Test}$	df	sign.?	Hypothese bestätigt?
V,2-GR	DT > DM ( $H_1$ )	1.08	1.33	28	nein	nein
V,2-BR	DT > DM ( $H_1$ )	1.07	1.29	28	nein	nein
V,2-RZ	DT > DM ( $H_1$ )	0.72	0.14	28	nein	nein
V,2-SZ <sup>a</sup>	DT=DM ( $H_0$ )	----	0.10	28	----	(nicht geprüft)
V,2-GZ <sup>a</sup>	DT=DM ( $H_0$ )	----	21.17	28	----	(nicht geprüft)

*Anmerkungen:*

GR: Anzahl gefundene Regeln; BR: Anzahl richtig benannter Regeln; RZ: relative Anzahl richtiger Zuordnungen; SZ: relative Anzahl sinnloser Zuordnungen; GZ: Gesamtanzahl der Zuordnungen;  $\alpha = .05$ . – <sup>a</sup> Die betreffenden Hypothesen werden nicht geprüft, da die als Voraussetzung für die inhaltlich sinnvolle Interpretation dieser Vergleiche, nämlich vorhersagekonforme Vortest-Nachtest-Veränderungen, nicht eingetreten sind (vgl. Text). – Wegen der Beibehaltung der vorhersagekonträren gerichteten und der vorhersagekonformen ungerichteten Nullhypothesen sowie der kleinen t-Werte wurden keine Effektgrößen  $d_{emp}$  berechnet (vgl. zur Begründung HAGER, 1987).

Daß in Anbetracht der auch durch den Ausfall einiger Kinder beim Nachtest geringen Stichprobengrößen und der daraus resultierenden geringen Teststärke die auf die KBA bezogenen Interaktionskontraste nicht die Grenze der statistischen Signifikanz erreichten, sollte nicht überbewertet werden, zumal wenn man bedenkt, daß sich teilweise bemerkenswerte differentielle Unterschiede bei den Vortest-Nachtest-Kontrasten unter den drei Programmen ergaben (vgl. Tab. 6). Es scheint uns daher berechtigt, diese Befundmuster trotz fehlender statistischer Signifikanz der Interaktionskontraste näher zu betrachten und zu interpretieren.

Erinnern wir uns: Bei der gewählten psychometrischen Erfassung des induktiven Denkens anhand der Subtests 3, 4 und 5 des CFT 1 zeigten sich für die beiden Denkfördergruppen und die Kontrollgruppe signifikante Leistungsanstiege. Das Bild ändert sich auch nicht, wenn man die Anstiege in den einzelnen Subtests getrennt betrachten würde (vgl. HAGER u. HÜBNER 1996). Anders sieht es jedoch bei der Untersuchung der induktiven Denkleistung mit Hilfe der Konzeptbildungsaufgaben aus. Hier ergeben sich lediglich für die Leistungsanstiege in den AVn

„Anzahl gefundener Regeln“ (GR) und „Anzahl richtig benannter Regeln“ (BR) unter allen drei Versuchsbedingungen signifikante Anstiege. Hinsichtlich der „relativen Anzahl der richtigen Zuordnungen“ (RZ), „relativen Anzahl der sinnlosen Zuordnungen“ (SZ) sowie der „Gesamtzahl der Zuordnungen“ (GZ) jedoch ergaben sich differentielle Ergebnisse. So ist für beide Denktrainingsgruppen ein Anstieg richtiger Zuordnungen nachzuweisen, der unter dem Kontrollprogramm ausbleibt. D.h. das Auffinden der Zuordnungsregel und deren Anwendung bei einwertigen Konzepten gelingt den Kindern, die mit einem der beiden Denkförderprogramme trainiert wurden, besser nach Abschluß der gezielten Denkförderung als nach Abschluß der Gedächtnisförderung.

Als Maß für Strategieveränderungen kann insbesondere die „Anzahl der sinnlosen Zuordnungen“ (SZ) bei den Konzeptbildungsaufgaben herangezogen werden. Entgegen unserer Erwartung zeigte sich ein signifikanter Abfall sinnloser Zuordnungen nach der Intervention jedoch nur für das Denktraining, nicht aber für das DenkMit, bei dem zudem die Vortest-Nachtest-Korrelation  $r_{VN}$  für die SZ mit .13 sehr gering ausfiel. Erwartungsgemäß stiegen die SZ in der Kontrollgruppe nicht an, und die dort auftretende Nullkorrelation ist mit diesem Programm ja durchaus vereinbar. Dieses Ergebnismuster kann als Hinweis darauf verstanden werden, daß die von KLAUER für sein Programm postulierte (aber aus den oben genannten Gründen bisher nicht nachgewiesene) Strategieverbesserung tatsächlich stattfindet. Das Ergebnis ist um so bemerkenswerter, wenn man berücksichtigt, daß in der vorliegenden Untersuchung trotz strikter Randomisierung der Kontrollgruppe die offensichtlich im Mittel leistungsstärkeren Kinder zugewiesen wurden (vgl. die IQ-Werte in Tab. 1).

Da die Subtests 3, 4 und 5 des CFT 1 weiter oben als „strategieanfällig“ und daher durch Zuwendungswirkungen beeinflußbar klassifiziert wurden, sollte man erwarten, daß dies auch für die (relative) Anzahl von sinnlosen Zuordnungen gilt. Dies scheint aber nicht der Fall zu sein, wie das Konstantbleiben unter dem DenkMit und unter dem Kontrollprogramm anzeigen. Dies könnte damit zusammenhängen, daß die Bearbeitung der Konzeptbildungsaufgaben individuell weniger strukturiert erfolgt als die Bearbeitung eines psychometrischen Tests, so daß *allgemeine* Strategieverbesserungen kaum wirksam werden können. Aber dies ist eine Post-hoc-Vermutung.

Bemerkenswert ist auch der Unterschied zwischen den Versuchsgruppen hinsichtlich der Gesamtanzahl der Zuordnungen (GZ; vgl. Tab. 7). Denn das statistisch signifikante Absinken dieser Zuordnungen beim DenkMit weist gegenüber der Konstanz der Anzahl der Zuordnungen beim Kontrollprogramm darauf hin, daß die mit dem DenkMit geschulten Kinder nach der Intervention reflexiver, also im Sinne einer Strategieverbesserung, vorgingen. Da eine reflexivere Strategie (Absinken der GZ) ja keineswegs bedeuten muß, daß auch weniger Fehler als vorher gemacht werden (vgl. auch LAUTH 1988, S. 154–156), spricht es für das DenkMit, daß mit dem Absinken der GZ auch Leistungsverbesserungen bei den richtigen Zuordnungen (RZ) einhergingen, wie bereits oben hervorgehoben wurde; entsprechende Leistungsverbesserungen zeigten sich auch bei den sinnlosen Zuordnungen (SZ), die jedoch die Signifikanzgrenze knapp verfehlten.

Beim Denktraining zeigten sich entsprechende Ergebnisse, allerdings mit dem Unterschied, daß hier vom Vor- zum Nachtest signifikant weniger sinnlose Zuordnungen (SZ) resultierten, während die Gesamtzahl der Zuordnungen (GZ) zwar absank, aber

nur innerhalb von noch durch den Zufall erklärbaren Grenzen. Anders ausgedrückt steht bei diesem Programm einer Verringerung der sinnlosen Zuordnungen ein Konstantbleiben der Gesamtzahl der Zuordnungen gegenüber, während sich dies beim DenkMit umgekehrt verhält.

Insgesamt deuten also die Befunde darauf hin, daß unter beiden Denkförderprogrammen eine Verbesserung der Bearbeitungsstrategien für induktive Anforderungen stellende Aufgaben erzielt worden ist, die allein durch unspezifische Interventionswirkungen nicht zu erklären ist, durch Wirkungen also, die darauf zurückzuführen sind, daß überhaupt trainiert wurde – wie etwa die Zuwendung. Besonders in Anbetracht des Umstandes, daß unter den beiden Denkförderprogrammen differentielle Verbesserungen resultierten, scheint es sich tatsächlich um „spezifische“ Programmwirkungen zu handeln, da eine Verbesserung der Motivation bzw. allgemeiner Bearbeitungsstrategien im Sinne OERTERS (1973) gerade wegen des höheren Ausgangsniveaus auch für die Kontrollgruppe nachweisbar hätte sein müssen.

## 6 Diskussion und Zusammenfassung

Wir werten diese Befunde insgesamt als Hinweis darauf, daß erstens programmbedingte Strategieverbesserungen aufgetreten sind und daß diese zweitens mit den üblichen psychometrischen Verfahren nicht nachweisbar sind, da diese nur Ergebnisse erfassen können, aber nicht Strategieveränderungen.

Die Wirkungen der untersuchten Förderprogramme auf den psychometrischen Wahrnehmungstest POD widersprechen in beiden Fällen den von den Trainingsautoren angestrebten Effekten. Während das DenkMit für sich neben der Denk- eine Wahrnehmungsförderung beansprucht, führte es im Vergleich zum Kontrollprogramm nicht zu einer Leistungssteigerung in diesem Bereich. Dagegen erwartet der Autor des Denktrainings keine besondere Wirkung seines Trainings auf Wahrnehmungsleistungen. Das Denktraining führt jedoch im Vergleich zum Kontrolltraining zu einer deutlichen Verbesserung in den Wahrnehmungsaufgaben des POD. Dieses Ergebnis bestätigt die Wahrnehmungshypothese von HAGER und HASSELHORN (1993b), die eine positive Wirkung des Denktrainings auf Wahrnehmungsleistungen postulieren.

Obwohl in bezug auf die psychometrischen AVn im Grunde die Befunde von HAGER und HASSELHORN (1995c) repliziert werden, muß es doch überraschen, daß die Auswirkungen der Denkförderprogramme auf die psychometrische AV CFT 1 die Auswirkungen eines Programms nicht überstiegen, das gerade auf den von den Denkförderprogrammen fokussierten Bereich nicht abzielt, denn auf der anderen Seite scheinen sich diese Befunde auf den ersten Blick nicht mit denen der Programmautoren zu decken. Doch die Wirkungen, über die SYDOW und MEINCKE (1994) für ihr DenkMit berichten, beruhen im wesentlichen auf sehr programmnahen AVn, so daß der Verdacht naheliegt, daß bei ihnen die Programmwirkungen artifiziell überschätzt werden, weil Coaching-Wirkungen in sie eingehen. Und die großen Effekte, die KLAUER (1989a, 1991) im Zusammenhang mit seinem Denktraining für Kinder I festgestellt hat, sind möglicherweise dadurch artifiziell erhöht, daß sie stets im Vergleich zu einer nicht-trainierten Versuchsgruppe, einer „Warte-



gruppe“, berechnet wurden. In einer Wartegruppe können aber keine Wirkungen der Tatsache auftreten, *daß* überhaupt interveniert wird. Aber üblicherweise werden diese und verwandte Wirkungen als so wichtig angesehen, daß viel Mühe darauf verwendet wird, sie zu kontrollieren (HAGER u. HASSELHORN 1995c; PERREZ et al. 1993; SCHMIDT-DENTER 1987, bes. S. 846), und STERNBERG und BHANA (1986, S. 61) konstatieren: „Most studies involved inadequate (untrained) control groups ...“. Auf jeden Fall blieb auch im vorliegenden Experiment eine ganz anders trainierte Versuchsgruppe bei induktiven Anforderungen stellenden AVn nicht hinter denjenigen Programmen zurück, die explizit induktives Denken verbessern wollen. Als plausibelste, wenn auch sicher nicht einzig mögliche Erklärung für diesen Befund drängen sich einmal mehr die (unspezifischen) Wirkungen der interventionsgebundenen Zuwendung und Motivierung auf.

Daß dessen ungeachtet die Denkförderprogramme auch spezifische Wirkungen erzeugen, zeigen die Ergebnisse bei den Konzeptbildungsaufgaben (KBA). Sie sind jedoch nicht sehr ausgeprägt; allerdings waren die entsprechenden Interaktionstests auch nicht sehr teststark. Es mag auch damit zusammenhängen, daß insbesondere beim Denktraining wenig bis keine Vorkehrungen getroffen werden, den Transfer des Geübten auf andere und dabei noch hinreichend ähnliche Aufgaben sicherzustellen. Denn dieser stellt sich – zumindest bei Kindern – nicht automatisch ein, sondern muß ebenfalls geübt werden (vgl. auch RESING 1996). Deshalb dürfte es Sinn machen, dieses Programm mit einem entsprechenden Transferbaustein anzureichern und gleichzeitig den Strategieaspekt stärker hervorzuheben; daran arbeiten wir derzeit.

## Literatur

- BECKER, P./SCHALLER, S./SCHMIDTKE, A. (1980): CPM – Coloured Progressive Matrices (2. Aufl.). Weinheim: Beltz.
- BRANDTSTÄDTER, J. (1982): Prävention als psychologische Aufgabe. In: BRANDTSTÄDTER, J./v. EYE, A. (Hrsg.): Psychologische Prävention. Grundlagen, Programme, Methoden. Bern: Huber, S. 15-36.
- BRUNER, J.S./GOODNOW, J.J./AUSTIN, G.A. (1956): A study of thinking. New York: Wiley.
- BÜCHEL, F.P./SCHARNHORST, U. (1993): Training des induktiven Denkens bei Lern- und Geistigbehinderten. In: KLAUER, K.J. (Hrsg.): Kognitives Training. Göttingen: Hogrefe, S. 95-123.
- ELSNER, B./HAGER, W. (1995): Ist das Wahrnehmungstraining von M. Frostig effektiv oder nicht? Zur Evaluation des Programmes zur Förderung der visuellen Wahrnehmung im deutschen Sprachraum. Praxis der Kinderpsychologie und Kinderpsychiatrie 44, 48-61.
- GRANT, D.A./BERG, E.A. (1948): A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. Journal of Experimental Psychology 38, 404-411.
- HAGER, W. (1987): Grundlagen einer Versuchsplanung zur Prüfung empirischer Hypothesen in der Psychologie. In: LÜER, G. (Hrsg.): Allgemeine Experimentelle Psychologie. Stuttgart: Fischer, S. 43-264.
- HAGER, W. (1995): Planung und Durchführung der Evaluation von kognitiven Förderprogrammen. In: HAGER, W. (Hrsg.): Programme zur Förderung des Denkens bei Kindern. Konzeption, Evaluation und Metaevaluation. Göttingen: Hogrefe, S. 100-206.
- HAGER, W./ELSNER, B. (1995): Validität einiger in der Trainingsforschung eingesetzter Tests. In: HAGER, W. (Hrsg.): Programme zur Förderung des Denkens bei Kindern. Konzeption, Evaluation und Metaevaluation. Göttingen: Hogrefe, S. 229-256.
- HAGER, W./ELSNER, B. (1996): Übung, Training, Coaching und Transfer – Die Rolle der Ähnlichkeit von Programm- und Kriteriumsaufgaben. In: KITTLER, U. et al. (Hrsg.): Bericht über das 2. Dortmunder Symposium für Pädagogische Psychologie (Arbeitstitel). Essen: Die Blaue Eule.

- HAGER, W./HASSELHORN, M. (1993a): Evaluation von Trainingsmaßnahmen am Beispiel von Klauers Denktraining für Kinder. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 25, 307-321.
- HAGER, W./HASSELHORN, M. (1993b): Induktives Denken oder elementares Wahrnehmen? Prüfung von Hypothesen über die Art der Wirkung eines Denktrainings für Kinder. *Empirische Pädagogik* 7, 421-458.
- HAGER, W./HASSELHORN, M. (1994): Kontroll- und Alternativtrainings bei der Evaluation von Trainingsprogrammen – und Retesteffekte. *Zeitschrift für Gerontopsychologie & -psychiatrie* 7, 169-177.
- HAGER, W./HASSELHORN, M. (1995): Zuwendung als Faktor der Wirksamkeit kognitiver Trainings. *Zeitschrift für Pädagogische Psychologie* 9, 163-179.
- HAGER, W./HÜBNER, S. (1996): Vergleichende Evaluation zweier Denkförderprogramme: erste Ergebnisse. In U. KITTLER et al. (Hrsg.): Bericht über das 2. Dortmunder Symposium für Pädagogische Psychologie (Arbeitstitel). Essen: Die Blaue Eule.
- HAGER, W./ROICK, T./BARTHOLOMÄUS, K. (1994): Test des induktiven Denkens (TID). Unveröffentlichtes Manuskript, Institut für Psychologie der Georg-August-Universität Göttingen.
- HASSELHORN, M./HAGER, W. (1995): Neuere Programme zur Denkförderung bei Kindern: Wie effektiv sind sie im Vergleich zu herkömmlichen Wahrnehmungsübungen? *Psychologie in Erziehung und Unterricht* 41, 221-233.
- HASSELHORN, M./HAGER, W. (1996): Neuere Programme zur Denkförderung bei Kindern: Bewirken sie größere Kompetenzsteigerungen als herkömmliche Wahrnehmungsübungen? *Psychologie in Erziehung und Unterricht* 43, 169-181
- HEATON, R.K. (1981): Wisconsin Card Sorting Test. Odessa, FL: Psychological Assessment Resources.
- HERRMANN, T. (1984): Methoden als Problemlösungsmittel. In: ROTH, E. (Hrsg.): Sozialwissenschaftliche Methoden. München: Oldenbourg, S. 18-46.
- HORN, J.L./CATTELL, R.B. (1966): Refinement and test of the theory of fluid and crystallized ability intelligence. *Journal of Educational Psychology* 57, 253-270.
- KAIL, R./PELLEGRINO, J.W. (1989): Menschliche Intelligenz (2. Aufl.). Heidelberg: Spektrum der Wissenschaft.
- KLAUER, K.J. (1989a): Denktraining für Kinder I. Göttingen: Hogrefe.
- KLAUER, K.J. (1989b): Die Messung von Transferdistanzen. Ein Verfahren zur Bestimmung der Unähnlichkeit von Aufgabenanforderungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 21, 146-166.
- KLAUER, K.J. (1991): Denktraining für Kinder II. Göttingen: Hogrefe.
- KLAUER, K.J. (1993a): Denktraining für Jugendliche. Göttingen: Hogrefe.
- KLAUER, K.J. (1993b): Evaluation einer Evaluation: Stellungnahme zum Beitrag von HAGER und HASSELHORN. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 25, 322-327.
- KLAUER, K.J. (1995a): Antworten und neue Befunde. *Zeitschrift für Pädagogische Psychologie* 9, S. 13-23.
- KLAUER, K.J. (1995b): Viel Lärm um nichts? Ein Kommentar. *Psychologie in Erziehung und Unterricht* 42, 295-299.
- KLAUER, K.J. (1995c): Weitere Erprobung des „Denktrainings für Jugendliche“ in der Oberstufe der Schule für Lernbehinderte. *Heilpädagogische Forschung* 21, 157-170.
- KLAUER, K.J./PHYE, G.D. (1994): Cognitive training for children. Seattle: Hogrefe & Huber.
- KLIX, F. (1983): Begabungsforschung – ein neuer Weg in der kognitiven Intelligenzdiagnostik? *Zeitschrift für Psychologie* 191, 360-387.
- KROWATSCHKE, D. (1994): Das Marburger Konzentrationstraining. Dortmund: Modernes Lernen.
- LAUTH, G.W. (1988): Trainingsmanual zur Vermittlung kognitiver Fertigkeiten bei retardierten Kindern. Tübingen: dgvt.
- LOCKOWANDT, O. (1993): Frostigs Entwicklungstest der visuellen Wahrnehmung – FEW (7. Aufl.). Weinheim: Beltz.
- LÜER, G./SPADA, H. (1992): Denken und Problemlösen. In: SPADA, H. (Hrsg.): Lehrbuch Allgemeine Psychologie. Bern: Huber, S. 189-280.
- MASENDORF, F./MAIHACK, V. (1986): Intelligenztraining bei Behinderten. *Die Rehabilitation* 25, 116-122.
- MATJUGIN, I.J./ASKOTSCHENSKI, T.J./BONK, I.A. (1995): Tastgedächtnis (2. Aufl.). Dortmund: Modernes Lernen.
- MATJUGIN, I.J./TSCHAKABERIA, H. (1996): Visuelles Gedächtnis (2. Aufl.). Dortmund: Modernes Lernen.
- MERZ, F. (1992): Buchbesprechung zu KLAUER (1989): Denktraining für Kinder I. Ein Programm zur intellektuellen Förderung. *Zeitschrift für Pädagogische Psychologie* 6, 152-156.

- OERTER, R. (1973): Steuerungskomponenten bei kognitiven Prozessen im Bereich schulischen Lernens. In: FREY, K./LANG, M. (Hrsg.): Kognitionspsychologie und naturwissenschaftlicher Unterricht. Bern: Huber, S. 56-78.
- PATRY, J.-L. (1994): Der Beitrag der Situationsspezifität zur Entstehung und Verringerung der Kluft zwischen Theorie und Praxis. In: OLECHOWSKI, R./ROLLETT, B. (Hrsg.): Theorie und Praxis. Aspekte empirisch-pädagogischer Forschung – quantitative und qualitative Methoden. Frankfurt am Main: Lang, S. 46-52.
- PELLEGRINO, J.W./GLASER, R. (1980): Components of inductive reasoning. In: SNOW, R.E./FEDERICO, P.-A./MONTAGUE, W.E. (Eds.): Cognitive process analyses of aptitude (Aptitude, learning, and instruction; vol. 1). Hillsdale, NJ: Erlbaum, S. 177-217.
- PERLETH, C. (1991): (Begabungs-) Diagnostik bei Schülern mit Lernbehinderung. In: HELLER, K.A. (Hrsg.): Begabungsdiagnostik in der Schul- und Erziehungsberatung. Bern: Huber, S. 236-276.
- PERLETH, C. (1992): Strategienutzung, Metagedächtnis und intellektuelle Begabung. Querschnitts- und Interventionsstudien bei Grundschulern. Dissertation an der Ludwigs-Maximilians-Universität zu München.
- PERREZ, M./HUBER, G.L./GEISSLER, K.A. (1993): Psychologie der pädagogischen Interaktion. In: WEIDENMANN, B./KRAPP, A./HOFER, M./HUBER, G.L./MANDL, H. (Hrsg.): Pädagogische Psychologie (3. Aufl.). München: PVU, S. 361-445.
- RESING, W.C.M. (1996): Untersuchungen zur Lernfähigkeit bei Kindern: Die Auswirkungen eines metakognitiven induktiven Denktrainings. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 28, 33-53.
- ROSSI, P.H./FREEMAN, H.E./HOFMANN, G. (1988): Programm-Evaluation. Stuttgart: Enke.
- SAUTER, F.C. (1979): Prüfung optischer Differenzierungsleistungen (POD). Braunschweig: Westermann.
- SCHMIDT-DENTER, U. (1987): Kognitive und sprachliche Entwicklungsförderung im Vorschulalter. In: OERTER, R./MONTADA, L. (Hrsg.): Entwicklungspsychologie (2. Aufl.). München: PVU, S. 814-853.
- SCHMUDE, C./SYDOW, H. (1994): Wirkungen kognitiver Trainings im Vorschulalter. Vortrag auf dem 39. Kongress der Deutschen Gesellschaft für Psychologie am 26.9.94 in Hamburg.
- STERNBERG, R.J./BHANA, K. (1986): Synthesis of research on the effectiveness of intellectual skills programs: Snake-oil remedies or miracle cures? Educational Leadership 44(2), 60-67.
- STERNBERG, R.J./GARDNER, M.K. (1983): Unities in inductive reasoning. Journal of Experimental Psychology: General 112, 80-116.
- SYDOW, H./MEINCKE, J. (1994): DenkMit. Das Berliner Programm zur Förderung des Denkens und der Wahrnehmung von drei- bis sechsjährigen Kindern. Kirchdorf: ZAK.
- WEISS, R.H./OSTERLAND, J. (1980): Grundintelligenztest CFT 1 – Skala 1 (4. Aufl.). Braunschweig: Westermann.

Anschrift der Verfasser/in: Prof. Dr. Willi Hager; Dipl.-Psych. Sigrid Hübner, Institut für Psychologie der Universität Göttingen, Goßlerstr. 14, 37073 Göttingen.