

Empirische Sonderpädagogik, 2012, Nr. 3/4, S. 331–347

## Untersuchungen zum Einfluss der Testmotivation auf die Testergebnisse des FEW-2

Peggy Beuthan, Willi Geser<sup>1</sup>, Claudia Schusterschitz<sup>2</sup>

<sup>1</sup>Leopold-Franzens Universität Innsbruck

<sup>2</sup>University for Health Sciences, Medical Informatics and Technology, Austria

### Zusammenfassung

Wenn Kinder in der ergotherapeutischen Diagnostik Aufgaben mit mangelhafter Testmotivation bearbeiten, können die Testergebnisse verfälscht werden. Das Ziel dieser Arbeit war es, zu untersuchen, ob ein Zusammenhang zwischen der Testmotivation und den Testergebnissen des FEW-2 besteht und ob sich die Testmotivation durch die Ankündigung einer Belohnung verbessert. Im Rahmen einer experimentellen Feldstudie wurden bei gesunden, deutschsprachigen Kindern ( $N = 80$ ) im Alter von 4 bis 6 Jahren die Testmotivation mit dem NV-MSVT vor (t1) und nach der (t2) Durchführung des FEW-2 evaluiert. Die Korrelationsanalysen ergaben, dass die Testmotivation die Testergebnisse im FEW-2 bei erhöhtem Anspruch an die kognitiven Leistungen (z. B. Konzentrationsleistung) beeinflusst. Die Ergebnisse der t-Tests belegten, dass eine angekündigte Belohnung keinen positiven Einfluss auf die Testmotivation ausübte. Der nachgewiesene Effekt der Testmotivation auf die Ergebnisse im FEW-2 unterstreicht die Notwendigkeit zur Überprüfung der Testmotivation für eine zuverlässige und valide neuropsychologische Diagnostik.

Schlüsselwörter: Diagnostik, Pädiatrie, Testmotivation, FEW-2

### Studies on the impact of achievement motivation on childrens' performance in the FEW-2

#### Abstract

If children in occupational diagnostics solve tasks with suboptimal test motivation, the results of such tests can be falsified. Against that background, the present research aims were to examine, whether there is an influence of children's test motivation on their performance in the FEW-2, and if their test motivation can be improved by promising a reward. In an experimental field study of 80 healthy and German-speaking children, at the age of 4 to 6 years, data on test motivation were collected and the FEW-2 was performed. To answer the research questions correlation analysis and t-tests for unpaired samples were carried out. The correlation analyses showed an influence of the motivation of the test on the outcomes of the FEW-2, if the claim of other cognitive services, such as, for example, the concentration of power was increased. According to the findings of the t-tests, the reward did not lead to a heightened test motivation amongst the participants of the experimental group. The revealed influence of test motivation on the outcomes of the FEW-2 underlines the necessity to test for test motivation in view of a reliable and valid neuropsychological diagnostic.

Keywords: Diagnostic, pediatrics, test motivation, FEW-2

## Einleitung

Bei der Durchführung neuropsychologischer Leistungstests (z. B. zur Überprüfung visueller Wahrnehmungsleistungen) im Rahmen der Befunderhebung vor einer Ergotherapie ist eine ausreichende Motivation der Kinder bei der Bearbeitung der gestellten Aufgaben für die Validität der Testergebnisse unverzichtbar (Merten & Brockhaus, 2004). Beim Frostigs Entwicklungstest der visuellen Wahrnehmung 2 (FEW-2) in seiner deutschsprachigen Fassung von Büttner, Dacheneder, Schneider und Weyer (2008) kommt es aufgrund der Dauer des Verfahrens (30 bis 45 Minuten) insbesondere bei Mädchen und Jungen im Alter von vier bis sechs Jahren häufig zu Motivationsproblemen. Diese äußern sich z. B. in Fragen der Kinder, ob sie alle Aufgaben lösen müssten oder wann die Erhebung beendet sei. Andere Mädchen und Jungen schildern offen ihre Unlust oder sie entscheiden sich für eine beliebige Lösung, ohne sich die Aufgabe anzusehen. Bei der „Symptomvalidierungstestung“ (Rheinberg, 2004) handelt es sich um ein Verfahren zur Beurteilung der aktuellen Motivation und zur Erkennung von suboptimalem Leistungsverhalten. Anders als in vielen anderen Ländern, in denen solche objektiven Testverfahren bei Erwachsenen und Schulkindern routinemäßig angewandt werden (Merten & Brockhaus, 2004), sind Symptomvalidierungstests in Deutschland aktuell wenig verbreitet (Merten & Brockhaus, 2004). Dabei wäre laut Iverson (2006) der Einsatz objektiver Instrumente wünschenswert: „Neuropsychologists should routinely assess for poor effort during testing and exaggerated symptoms and problems“ (S. 83). Dies erscheint ganz besonders dringlich, da Büttner, Dacheneder, Schneider und Weyer (2008) bereits auf einen möglichen Einfluss der Motivation auf die Testergebnisse des FEW-2 hingewiesen haben.

Im therapeutischen Alltag werden verschiedene Belohnungsstrategien zur Ver-

besserung der Testmotivation eingesetzt. Auch wenn eine Belohnung nicht zwingend bei jedem Kind eine motivationsfördernde Wirkung nach sich zieht, ist davon auszugehen, dass die Motivation nicht negativ beeinflusst wird (Heise, Sauep & Johanning, 2008). Die Sicherstellung einer ausreichenden Motivation der Kinder beim FEW-2 ist notwendig, um (1) eine verlässliche und realitätsnahe Befunderhebung im Rahmen der ergotherapeutischen Diagnostik durchzuführen, um (2) eine zuverlässige Früherkennung visueller Wahrnehmungsstörungen zu ermöglichen und um (3) die Entstehung anderer Entwicklungsstörungen als Folge einer visuellen Wahrnehmungsstörung durch eine rechtzeitige Behandlung zu vermeiden. Darüber hinaus dient die Überprüfung der Testmotivation während neuropsychologischer Verfahren der Qualitätssicherung der Arbeit der Therapeutin bzw. des Therapeuten und verhindert die Festlegung fehlerhafter Normwerte (Merten & Brockhaus, 2004).

## Testmotivation

Die im FEW-2 erbrachte Leistung ist das Ergebnis des Zusammenspiels von Vorkenntnissen, der Geschwindigkeit der Informationsverarbeitung und der Testmotivation (Baumert & Demmrich, 2001). Während in manchen Testsituationen unzureichende spezifische Vorkenntnisse durch die allgemeine Intelligenz ausgeglichen werden können (Baumert & Demmrich, 2001), ist die Testmotivation nicht kompensierbar (Wang, Haertel & Walberg, 1993). Sie wird in diesem Zusammenhang als die Bereitschaft definiert, den Aufwand und die Anstrengung zur Lösung der Testaufgaben aufzubringen (Baumert & Demmrich, 2001). Nach der Erwartungswert-Theorie ist die Motivation als eine Form der Leistungsmotivation zu verstehen (Rheinberg, 1995). Die Ebene der Motivation ist jedoch eine Funktion der erwarteten Leistung

und bestimmt sich durch die persönliche Bedeutung, die eine getestete Person der Bedeutung der diagnostischen Situation beimisst (Baumert & Demmrich, 2001). Der individuelle Wert eines Tests bestimmt sich für eine Person einerseits durch seine direkten oder indirekten Konsequenzen, andererseits aber auch durch die anregende Erfahrung, an einer schwierigen Aufgabe beteiligt zu sein, durch die Zufriedenheit, die Aufgabe bewältigt zu haben, oder durch Fähigkeitsinformationen, welche die Person durch den Test erhält (Baumert & Demmrich, 2001). Des Weiteren ist die persönliche Bedeutung der Aufgabe nicht nur von der motivationalen Einstellung des Individuums abhängig, sondern auch von den Eigenschaften der Aufgabe selbst, dem sozialen Kontext der Person und den sich daraus ergebenden Konsequenzen. Entscheidenden Einfluss auf die Motivation haben intrinsische Werte, wie die mit der Fähigkeitsinformation in Zusammenhang stehenden Verwirklichungs- und Nützlichkeitswerte, die mit den direkten oder indirekten Konsequenzen der Leistung verbunden sind (Wigfield & Eccles, 1992). In Bezug auf den FEW-2 bedeuten diese Überlegungen, dass die Testmotivation der Kinder davon abhängt, ob sie intrinsisch dazu bereit sind, das Verfahren durchzuführen, ob sie einen Nutzen darin sehen, den Aufwand und die Anstrengung aufzubringen, und ob sie glauben, die Aufgaben lösen zu können. Des Weiteren haben Baumert und Demmrich (2001) nachgewiesen, dass der Einfluss der Nützlichkeitskomponente auf die Testausführung von der Konzentration und der Lust abhängig ist, das Instrument bis zum Ende zu bearbeiten.

Petermann und Macha (2005) haben Empfehlungen zur Optimierung der Motivation für psychologische Testverfahren bei Kindern formuliert. Demnach sind angemessene äußere Umstände zu gewährleisten, die den Bedürfnissen und Ansprüchen der Mädchen und Jungen gerecht werden

(z. B. physiologische Arbeitshaltung, ausreichende Beleuchtung, gute Belüftung). Besonders bei jüngeren Kindern kann eine ungünstige Umgebung (z. B. offene Schränke und Regale, viele Bilder an den Wänden, erhöhter Geräuschpegel) erheblichen Einfluss auf die Konzentration und Motivation ausüben. Um von Beginn an eine „freundliche Atmosphäre“ herzustellen, ist es notwendig, dass die physiologischen Bedürfnisse der Mädchen und Jungen (z. B. Essen, Trinken) vorher befriedigt werden und dass die Diagnostikerin bzw. der Diagnostiker mit dem Testverfahren vertraut ist. Unsichere Instruktionen oder das Nachlesen von Bewertungsvorschriften unterbrechen den Kontakt zum Kind, was einen reibungslosen Ablauf der Testdurchführung gefährden kann. Grundsätzlich sollte angestrebt werden, die Kinder ohne Begleitpersonen zu testen, da Eltern sich häufig nicht angemessen passiv verhalten und verbal oder nonverbal versuchen, auf ihre Tochter bzw. ihren Sohn einzuwirken. Diese Interventionen können zu einem massiven Leistungsdruck für die Mädchen und Jungen führen, was im Hinblick auf valide Testergebnisse unbedingt zu verhindern ist (Petermann & Macha, 2005). Bei sehr ängstlichen oder unsicher gebundenen Kindern ist die Anwesenheit eines Elternteils unter Umständen jedoch nicht vermeidbar. In einem solchen Fall sollten die Eltern im Vorfeld entsprechend instruiert werden.

Für eine gute Testmotivation ist eine positive Interaktion während der Durchführung des Verfahrens unverzichtbar. Bei Leistungstests wie dem FEW-2 sollten keine Rückmeldungen zur Qualität der Aufgabenlösungen gegeben werden. Es ist lediglich die gute Mitarbeit zu loben. Auch bei falschen Antworten kann die Therapeutin bzw. der Therapeut freundlich und bekräftigend auf die Kinder eingehen. Droht die Testperson an der Vollendung einer Aufgabe zu scheitern, wirken sich leichte Hilfestellungen, die es ermöglichen, die Aufgabe

zu beenden, günstig auf den weiteren Motivationsverlauf aus. Grundsätzlich ist zu empfehlen, der Testsituation den Charakter einer Spielsituation zu verleihen, Instruktionen so zu gestalten, dass sie höflich zur Aufgabenbearbeitung auffordern, und (sofern es möglich ist) Handlungsspielräume zu schaffen, um die Testsituation offener zu gestalten (Petermann & Macha, 2005). Bei der Durchführung des FEW-2 ist es generell möglich, diese Empfehlungen zu berücksichtigen. Lediglich der Gestaltung einer offeneren Testsituation sind Grenzen gesetzt. Im Manual des Verfahrens wird nur für die vier- und fünfjährigen Kinder eine kurze Pause vorgeschrieben. Jedoch ist aufgrund der Form (Papier-Bleistift-Tests) und der Dauer des FEW-2 davon auszugehen, dass auch bei den sechsjährigen Kindern eine Unterbrechung dem Bewegungsdrang gut täte. Im Manual wird nicht ausdrücklich davon abgeraten.

### **Der Einsatz von Belohnung zum Abbau von Motivationsproblemen**

Die Motivationsform hängt von der persönlichen Bedeutung des Verfahrens für das zu testende Individuum ab (Baumert & Demmrich, 2001). Die Identifikation der individuellen Einflussfaktoren der motivationalen Ebene und die Kenntnis von Strategien, wie man eine optimale Motivation erzielen kann, sind für die erfolgreiche Durchführung des FEW-2 hilfreich.

Rheinberg (2004) unterscheidet vier Motivationsformen. Die erste Motivationsform wird als Selbstinitiative (spontane Aktivität) bezeichnet und umschreibt einen Prozess, bei dem die Tätigkeit selbst schon ansprechend ist und Spaß bereitet. Nach Heckhausen (1991) wäre die bzw. der zu Testende unter diesen Bedingungen intrinsisch motiviert, das Verfahren durchzuführen. Von einer fremd kontrollierten Aktivität (zweite Motivationsform) spricht man, wenn eine Tätigkeit keinen Spaß

macht, aber trotzdem ausgeführt wird, weil es bedeutsame Personen erwarten. Kann die zu testende Person nicht in dem Maße motiviert werden, dass sie eine Aufgabe fremd kontrolliert ausübt, können Motivationsprobleme auftreten (Rheinberg, 2004). Führt die Aktivität zu keinem Ergebnis, besteht ein vollständiges Motivationsdefizit. Nach Heckhausen (1991) sieht die bzw. der zu Testende dann keinen Nutzen in der Tätigkeit. Trotz eines zu erwartenden Ergebnisses kann beim Fehlen von lohnenden Folgen (Rheinberg, 1995) ein Anreizdefizit bestehen. Wenn bei der Testperson Zweifel an der Erreichbarkeit des sich lohnenden Ergebnisses aufkommen, dann liegt ein Wirksamkeitsdefizit vor (Rheinberg, 2004). Das zu testende Individuum ist sich unsicher, ob es das Ziel der Aktivität durch sein Wirken erreichen kann (Heckhausen, 1991). Voraussetzungen für eine selbst gesteuerte Zielaktivität (dritte Motivationsform) sind das Wissen, dass die eigene Aktivität zu einem lohnenden Ergebnis führen kann und dass die Tätigkeit keinen Verzicht auf eine andere interessante Tätigkeit fordert (und damit keine Ablehnung erzeugt). Wird eine Aufgabe aus einem der oben genannten Gründe abgelehnt, kann bei unzureichender Selbstregulationskompetenz ein Volitionsdefizit als Motivationsproblem auftreten (Rheinberg, 2004). Besonders bei langweiligen und anstrengenden Tätigkeiten, bei ablenkenden und unangenehmen Durchführungsbedingungen oder bei unüberschaubar komplexen und schwierigen Tätigkeiten ist eine hinreichende Selbstregulationskompetenz unverzichtbar (Kuhl, 1996). Bei der vierten Motivationsform, der selbstbeherrschten Zielaktivität (Rheinberg, 2004), werden Aufgaben, die zwar mit Abneigung besetzt sind, unter volitionaler Anstrengung (Selbstregulation) mit Blick auf sich lohnende Folgen erfolgreich absolviert (Sokolowski, 1993). Das Diagnoseschema in Abbildung 1 verdeutlicht die unter-

schiedlichen Motivationsprobleme (Rheinberg, 2004).

Die in dem Diagnoseschema dargestellten Motivationsformen spielen auch beim FEW-2 eine Rolle: Neben Kindern, denen die Aufgaben von Beginn an Spaß bereiten und die somit den Test von sich aus motiviert bearbeiten, gibt es möglicherweise auch solche, die sich trotz des fehlenden

Spaaßes auf die Aufgaben einlassen, weil man es von ihnen erwartet (fremd kontrollierte Aktivität). Sprechen Mädchen und Jungen nicht auf Erwartungen von außen an, sollte das Ergebnis des FEW-2 für das betreffende Kind ersichtlich sein, damit es ein Ziel verfolgen kann und einen Zweck in seiner Handlung sieht, der den Aufwand und die Anstrengung aus seiner Sicht rechtfertigt. Das Ziel und der Zweck des FEW-2 lassen sich manchen Mädchen und Jungen nur schwer vermitteln. Nach dem Diagnoseschema von Rheinberg (2004) besteht dann ein vollständiges Motivationsdefizit. In dieser Situation müssten zur Motivationsförderung freudvolle Handlungsbedingungen bzw. attraktive Tätigkeiten geboten, die Fremdkontrolle verstärkt oder erkennbare Zwischenergebnisse eingefügt werden. Abwandlungen des Testschemas sind jedoch unzulässig, so dass alternativ eine Belohnung das Ergebnis am Ende ersetzen könnte. Dabei muss die Anerkennung für das Kind individuell ansprechend sein, damit es die Motivation aufbringt, den Test durchzuführen. Ist das nicht der Fall, dann besteht ein Anreizdefizit. Bei Anreizdefiziten sind eng an das Ergebnis gekoppelte, attraktive Ergebnisfolgen einzuführen, um die Motivation aktuell zu steigern. Hat das Kind nach der Vorstellung des FEW-2 das Gefühl, dass

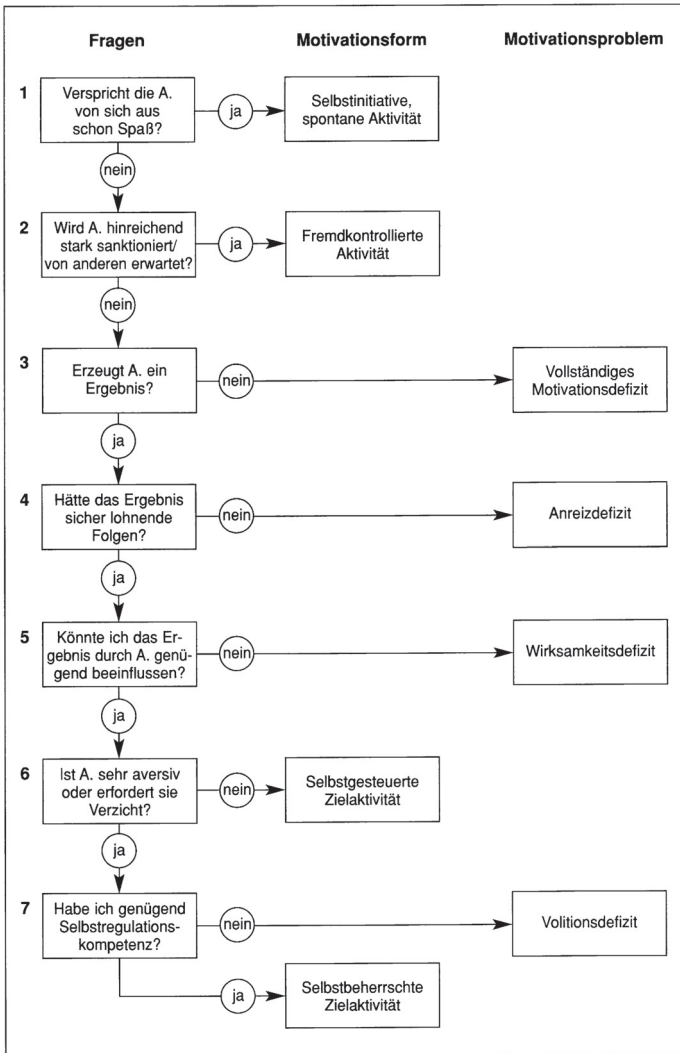


Abbildung 1: Diagnoseschema – Frage- und Antwortsequenz zur Bestimmung verschiedener Motivationsformen und -probleme (A = Aktivität) (Rheinberg, 2004).

Das Ziel und der Zweck des FEW-2 lassen sich manchen Mädchen und Jungen nur schwer vermitteln. Nach dem Diagnoseschema von Rheinberg (2004) besteht dann ein vollständiges Motivationsdefizit. In dieser Situation müssten zur Motivationsförderung freudvolle Handlungsbedingungen bzw. attraktive Tätigkeiten geboten, die Fremdkontrolle verstärkt oder erkennbare Zwischenergebnisse eingefügt werden. Abwandlungen des Testschemas sind jedoch unzulässig, so dass alternativ eine Belohnung das Ergebnis am Ende ersetzen könnte. Dabei muss die Anerkennung für das Kind individuell ansprechend sein, damit es die Motivation aufbringt, den Test durchzuführen. Ist das nicht der Fall, dann besteht ein Anreizdefizit. Bei Anreizdefiziten sind eng an das Ergebnis gekoppelte, attraktive Ergebnisfolgen einzuführen, um die Motivation aktuell zu steigern. Hat das Kind nach der Vorstellung des FEW-2 das Gefühl, dass

es die Aufgaben aufgrund ihrer Anzahl oder ihrer Komplexität nicht lösen kann, besteht ein Wirksamkeitsdefizit. Es ist jedoch möglich, dass ein solches Problem auch dann auftritt, wenn die Diagnostikerin bzw. der Diagnostiker das Testverfahren nicht vorstellt und direkt mit der Durchführung beginnt. Dies ist eventuell dann der Fall, wenn das Kind nicht abschätzen kann, ob seine Fähigkeiten zur Lösung der Aufgaben ausreichen. Zur Vermeidung von Wirksamkeitsdefiziten sollten Strategien eingeübt werden, die die Zuversicht der Mädchen und Jungen stärken, die Aufgaben lösen zu können. Aufgrund der eingeschränkten Variationsmöglichkeiten ist die einfachste Lösung aber eine individuell angepasste Vorstellung des FEW-2 seitens der Diagnostikerin bzw. des Diagnostikers, die ein Gefühl der Überforderung beim Kind vermeidet. Fehlen weitere Motivationshindernisse, vollzieht es eine selbst gesteuerte Zielaktivität. Problematisch für die Durchführung des FEW-2 können aber auch eine große Abneigung gegen die Aufgaben oder konkurrierende Bedürfnisse des betreffenden Mädchens bzw. Jungens sein. Abhängig von seiner Selbstregulationsfähigkeit kann das Kind trotzdem die Motivation aufbringen (selbstbeherrschte Zielaktivität), die Aufgaben durchzuführen. Oder es scheitert aufgrund von Volitionsdefiziten. Bei Volitionsdefiziten sollte versucht werden, die Abneigung gegen die Aufgaben und die Schwierigkeit bei der Zielverfolgung zu verringern (ebd.). Ob und wie ein Kind den Test bis zum Ende durchführt, hängt daher von vielen Faktoren ab.

### **Motivationsdiagnostik bei Simulationsverdacht**

Mit Hilfe der Symptomvalidierungstestung kann beurteilt werden, ob die Testperson eine gute Motivation zeigt und damit valide Ergebnisse produziert, oder ob eine Simulation vorliegt. In diesem Kontext wird

der Begriff der Simulation im Sinne eines suboptimalen Leistungsverhaltens verwendet (Merten, 2002). Dabei erbringt die Person in einem Test eine Leistung, die unterhalb ihrer tatsächlichen Fähigkeiten liegt (Merten, 2002). Suboptimales Leistungsverhalten tritt nicht nur durch bewusste Vortäuschung von Krankheitssymptomen auf, sondern kann auch durch ungünstige Umstände hervorgerufen werden (Merten, 2002).

Symptomvalidierungstests beurteilen die „Echtheit“ von Symptomen und Störungen sowie die Echtheit von neuropsychologischen Testergebnissen (Merten, 2002). Ein Symptomvalidierungstest muss auch für Testpersonen durchführbar sein, die an schweren neuropsychologischen Erkrankungen leiden (Merten & Brockhaus, 2004). Der Bereich des Gehirns, in dem das Gedächtnis zur Wiedererkennung (Recognition Memory) liegt, ist immun gegen eine Vielzahl von Schädigungen und Traumata. Die Wiedererkennung ist eine stabile Komponente der Gehirnleistung, da bereits junge Kinder einen hohen Leistungsstand in diesem Bereich aufweisen (Piaget & Inhelder, 1973; Brown, 1975; Perlmutter & Lange, 1978). Sophian und Stigler (1981) analysierten die Wiedererkennung von Gesichtern bei Kindern im Alter von ca. drei, fünf und sieben Jahren sowie bei College-Studierenden. Bei den Vorschulkindern gab es in den Altersgruppen keine Unterschiede bezüglich der Wiedererkennungsleistung; Verbesserungen traten lediglich von der Vorschulperiode bis zur ersten Klasse auf. Zahlreiche Autorinnen und Autoren stellen das Wiedererkennen als eine Leistung dar, die in jedem Alter messbar ist und die zugleich unverändert bleibt (Weiner, 1994). Damit ist diese Gedächtnisleistung ideal für die Symptomvalidierung im Rahmen neuropsychologischer Leistungstests von Kindern im Vorschulalter (Rohling, 2004). Bei den Symptomvalidierungstests werden Wörter, Nummern oder Symbole verwen-

det, die eine semantische Verbindung mit einem anderen Merkmal hervorrufen.

### **Zum Einfluss von Testmotivation auf neuropsychologische Leistungstests**

Green, Rohling, Lees-Haley und Allen (2001) untersuchten in einer experimentellen Studie den Zusammenhang von Testmotivation und neuropsychologischen Leistungstests. Die Ergebnisse der 904 Patientinnen und Patienten belegten deutlich, dass die Motivation, gemessen mit dem Word Memory Test (WMT) von Green, Allen und Astner (1996), die mit Abstand wichtigste Einflussvariable für die Ergebnisse verschiedener neuropsychologischer Verfahren darstellt: 50 % der Gesamtvarianz innerhalb dieser Tests konnten durch diese Variable aufgeklärt werden. Von den Patientinnen und Patienten bestanden 8 % nicht, was laut den Autoren auf ein suboptimales Leistungsverhalten bei den neuropsychologischen Leistungstests schließen ließ, so dass die Testergebnisse nicht ausgewertet werden konnten.

In einer weiteren Studie wurde an 69 Erwachsenen mit leichtem Schädel-Hirn-Trauma der Einfluss der Testmotivation auf die Ergebnisse der Halstead-Reitan Battery (HRB) von Reitan und Wolfson (1993) untersucht. Zur Messung der Testmotivation wurde der Test of Memory Malingering (TOMM) von Tombaugh (1996) verwendet. 47 % der Gesamtvarianz der Testergebnisse in der HRB konnte durch die Testmotivation geklärt werden (Constantinou, Bauer & Ashendorf, 2005).

In einer groß angelegten experimentellen Studie wurde an 1307 Probandinnen und Probanden der Einfluss der Testmotivation auf 23 verschiedene neuropsychologische Leistungstests belegt: "The pervasive influence of effort on almost all neuropsychological tests may be readily seen ..." (Green, 2007, S. 49). Zum Einfluss der Test-

motivation auf das Abschneiden im FEW-2 oder in anderen visuellen Wahrnehmungstests für Kinder wurden bisher noch keine Untersuchungen durchgeführt.

Auch Mädchen und Jungen im Vorschulalter sind schon in der Lage, Testergebnisse zu verfälschen. Zum einen sind Kinder in diesem Alter leicht durch Autoritätspersonen zu beeinflussen. Eltern könnten beispielsweise aus finanziellen Gründen ihre Töchter und Söhne zu Antworten animieren, die den eigenen Vorstellungen der Kinder widersprechen. Diesen Umstand beschreibt Rohling (2004) wie folgt:

"... even if a child's performance is considered to be intentionally poor, it is possible that a 'concerned' parent or a 'savvy' attorney with financial incentives was instrumental in coaching the child to respond to some questions incorrectly. Some children may believe that these authority figures' advice should be followed unchallenged, even if it runs contrary to their experience" (S. 21).

Zum anderen glauben Kinder häufig, dass ihnen mangelnde Anstrengungsbereitschaft oder Krankheit zu mehr Aufmerksamkeit und Liebe verhilft (Rohling, 2004). Auch Mädchen und Jungen im Vorschulalter sind bereits in der Lage, einige der aktuellen Testverfahren zur Überprüfung der Motivation zu bewältigen, auch wenn diese zunächst für Erwachsene konzipiert wurden. Constantinou und McCaffrey (2003) überprüften 128 gesunde Kinder im Alter von 5 bis 12 Jahren mit den Testverfahren TOMM und Rey-Fifteen Items Test (FIT) von Rey (1958). Nur zwei der Kinder erzielten im TOMM Ergebnisse unter den Cut-off-Werten für Erwachsene; es war kein Alterseffekt zu verzeichnen. Im FIT schnitten ältere Kinder dagegen besser ab als jüngere.

In einer Studie wurden 135 Kinder im Alter von 7 bis 18 Jahren mit dem WMT einer genauen Diagnostik unterzogen. Nur 19 Mädchen und Jungen wiesen Testwerte auf, die unter dem Cut-off-Point der

Erwachsenen lagen. Die Autoren schlussfolgerten, dass die Resultate in keinem Zusammenhang zum Alter und zur Intelligenz der Kinder stehen. Sobald Mädchen und Jungen eine Lesefertigkeit des dritten Grades erworben haben, ist es beim WMT möglich, die Cut-off-Werte für Erwachsene anzuwenden (Green & Flaro, 2003).

Courtney, Dinkins, Allen und Kuroski (2003) führten Untersuchungen an 111 lern- und verhaltensauffälligen Kindern im Alter von sechs bis 17 Jahren durch. Mädchen und Jungen unter zehn Jahren schnitten im WMT und im Computerized Assessment of Response Bias (CARB) von Allen, Conder, Green und Cox (1997) signifikant schlechter ab. Kinder ab 10 Jahren erzielten in den Tests vergleichbare Ergebnisse wie Erwachsene. Darüber hinaus war die Leistung abhängig von ihrer Lesefertigkeit. Der CARB wurde auch von Green verwendet, bis er 1994 den WMT entwickelte (Green, 2004). Da die Sensitivität der Verfahren zur Testmotivationsmessung bei geringem Alter, niedrigen Intelligenzlevel und geringer Lesefertigkeit variieren, ist bisher noch unklar, welches Instrument für sehr junge Kinder am besten geeignet ist (Rohling, 2004).

Im Rahmen dieser Arbeit soll der Einfluss der Testmotivation auf die Ergebnisse des FEW-2 und der Einfluss einer Belohnung auf die Testmotivation bzw. das Antwortverhalten von Kindern untersucht werden. Da der FEW-2 aufgrund seiner Dauer (30 bis 45 Minuten) einen erhöhten Anspruch an die Konzentration und Aufmerksamkeit der Kinder stellt, ist von einer abnehmenden Testmotivation im Testverlauf auszugehen. Daher sollte die Motivation zu zwei verschiedenen Zeitpunkten beurteilt werden.

## **Methode**

### ***Versuchspersonen***

Im Rahmen einer experimentellen Feldstudie wurden 80 gesunde, deutschsprachige

Kinder im Alter von vier bis sechs Jahren untersucht. Die Datenerhebungen erfolgten in Kindertagesstätten. Eine Zuteilung der Mädchen und Jungen zur Experimental- ( $N = 40$ ) bzw. zur Kontrollgruppe ( $N = 40$ ) erfolgte randomisiert.

### ***Instrumente***

Der FEW-2 besteht aus acht Subtests: (1) Auge-Hand-Koordination (Cronbach's Alpha = 0,87), (2) Raum-Lage-Wahrnehmung (Cronbach's Alpha = 0,88), (3) Abzeichnen (Cronbach's Alpha = 0,87), (4) Figur-Grund-Wahrnehmung (Cronbach's Alpha = 0,78), (5) Wahrnehmung räumlicher Beziehungen (Cronbach's Alpha = 0,81), (6) Gestaltschluss (Cronbach's Alpha = 0,84), (7) visuo-motorische Geschwindigkeit (Cronbach's Alpha = 0,79) und (8) Formkonstanz (Cronbach's Alpha = 0,80) (Büttner et al., 2008). Bei den vierjährigen Mädchen und Jungen wird der Subtest visuo-motorische Geschwindigkeit laut Manualvorgabe nicht durchgeführt und bei den vier- und fünfjährigen Kindern wird eine zusätzliche Pause von fünf Minuten nach dem Subtest Figur-Grund-Wahrnehmung eingelegt. In der Einzelsituation dauert das Verfahren zwischen 30 und 45 Minuten. Der Test fand vormittags zwischen 8 und 12 Uhr statt, da die Kinder hier voraussichtlich noch über eine relativ hohe Konzentration bzw. Aufmerksamkeit verfügen und noch nicht vom Tagesgeschehen erschöpft sind.

Zur Messung der Testmotivation wurde der NV-MSVT (Nonverbal Medical Symptom Validity Test) von Green (2008) ausgewählt. Dabei handelt es sich um einen standardisierten, computergestützten Symptomvalidierungstests zur Erkennung von suboptimalem Leistungsverhalten (Green, Gervais & Flaro, 2006). Der NV-MSVT besteht aus vier „nachweisbar einfachen“ Subtests (Immediate Recognition Trial [IR], Delayed Recognition Trial [DR], Delayed Recognition Trial/Archetypes [DRA], De-



layed Recognition Trial/ Variations [DRV]) und aus zwei „schwierigeren“ Subtests zur Gedächtnisleistung (Paired Associates [PA] und Free Recall [FR]; vgl. Testbodeneffekt in Merten, 2002).

### Datenerhebung

Vor der Testdurchführung wurde den Kindern der FEW-2 als Mal- und Zeigenspiel vorgestellt (Eingangsinstruktion). Sie erhielten dabei lediglich Aufgaben mit geringer Komplexität und niedriger Schwierigkeit (Petermann & Macha, 2005). Den Mädchen und Jungen war jedoch nicht bekannt, dass zusätzlich ein Test zur Symptomvalidierung durchgeführt wurde. Der Eingangsinstruktion folgte bei den Kindern der Experimentalgruppe die Inaussichtstellung einer Belohnung in Form einer „Goldmedaille“ für gute Anstrengungsbereitschaft („Wenn du dich so gut wie möglich anstrengst, dann bekommst du am Ende eine Goldmedaille.“). Nach der Belohnungsankündigung wurde der erste Symptomvalidierungstest durchgeführt (t1), danach der FEW-2 und dann der zweite Symptomvalidierungstest (t2). Am Ende erhielten die Kinder ihre Belohnung. Bei der Kontrollgruppe wurde vor der Durchführung der Testverfahren keine Belohnung angekündigt. Der weitere Verlauf entsprach der Vorgehensweise bei der Experimentalgruppe. Abschließend erhielten auch die Kinder der Kontrollgruppe die Belohnung. In der folgenden Abbildung 2 sind beide Untersuchungsverläufe noch einmal zum Vergleich dargestellt.

Die Testmotivation wurde in beiden Untersuchungsgruppen zu zwei Messzeitpunk-

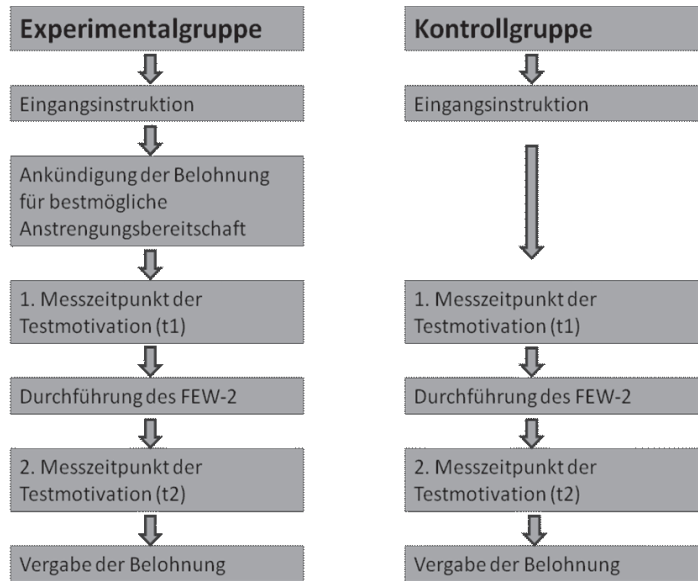


Abbildung 2: Darstellung der Untersuchungsverläufe mit (Experimentalgruppe) und ohne (Kontrollgruppe) Belohnung.

ten bestimmt, um Änderungen im Verlauf zu erkennen. Auch die in Aussicht gestellte Belohnung beeinflusste möglicherweise nicht nur das anfängliche Verhalten der Kinder, sondern auch die Selbstregulationsprozesse bei nachlassender Motivation im Zuge der Durchführung (Rheinberg, 2004).

### Auswertung

Während sich der FEW-2 entweder durch eine Diagnostikerin bzw. einen Diagnostiker oder mit Hilfe eines Computerprogramms auswerten lässt, werden die Ergebnisse des NV-MSVT automatisch von einem Computerprogramm analysiert. Zunächst berechnet es das Leistungsverhalten des getesteten Individuums. Die Person hat den Test bestanden, wenn der Cut-Off-Wert von mindestens 88 % erreicht wird  $(DR + CNS + DRA + DRV)/4$ . Die Consistency (CNS) ergibt sich aus der Differenz der Subtests IR und DR. Bei Ergebnissen unterhalb 88 % gilt der Symptomvalidierungstest als

nicht bestanden. Dann ermittelt das Computerprogramm weiter, ob die Testperson die Ergebnisse aufgrund mangelhafter Testmotivation verfälscht hat oder an Demenz leidet.

Die an Demenz erkrankten Testpersonen bewältigen die eher einfach konstruierten Motivationssubtests (IR, DR, DRA, DRV) gut, schneiden aber in den beiden letzten Subtests (PA und FR), die hauptsächlich Gedächtnisleistungen messen, schlecht ab. Bei Personen, welche die Testergebnisse aufgrund mangelhafter Motivation verfälschen, befinden sich die Leistungen aller Subtests unterhalb des Normbereiches. Außerdem liegen die Indices im PA deutlich höher, als die Ergebnisse im DR, DRA, DRV und in der CNS. Als Güteindizes wurden für den NV-MSVT eine Sensitivität von 0,60 und eine Spezifität von 0,98 im Hinblick auf suboptimales Leistungsverhalten ermittelt. Damit liegt die Wahrscheinlichkeit,

dass Testpersonen mit suboptimalem Leistungsverhalten im Test auffällig werden, bei 60%. Hingegen beträgt die Wahrscheinlichkeit, dass Testpersonen mit guter Testmotivation im Test gut abschneiden, 98% (Green, 2008). Um die Fragestellungen zu beantworten, erfolgte die Durchführung von Korrelationsanalysen und von t-Tests mithilfe von SPSS. Bei signifikanten Ergebnissen wurde sich an der 5%-Grenze der Irrtumswahrscheinlichkeit orientiert. Alle hochsignifikanten Ergebnisse entsprechen einem Signifikanzniveau von 1%.

## Ergebnisse

Zur Überprüfung der Annahme, dass die Testmotivation das Abschneiden im FEW-2 beeinflusst, wurden Korrelationsanalysen zwischen dem Gesamtscore des NV-MSVT und den einzelnen Subtests des FEW-2 getrennt für die Experimentalgruppe (mit

*Tabelle 1: Korrelationskoeffizienten zwischen den einzelnen FEW-2-Subtestergebnissen und dem Gesamtscore des NV-MSVT zum ersten Messzeitpunkt t1.*

Subtests FEW-2	Experimentalgruppe	Kontrollgruppe
	(mit Belohnung)	(ohne Belohnung)
Auge-Hand-Koordination	0,14	-0,17
Lage im Raum	0,40*	0,12
Abzeichnen	0,40*	0,29
Figur-Grund-Wahrnehmung	0,26	0,13
Räumliche Beziehungswahrnehmung	0,41*	0,31
Gestaltschluss	0,20	0,08
Visuo-motorische Geschwindigkeit	0,34	-0,16
Formkonstanz	0,30	0,28

*Legende: Signifikanz auf 5%-Niveau (\*), 1%-Niveau (\*\*)*

Belohnung) und die Kontrollgruppe (ohne Belohnung) durchgeführt. In der folgenden Tabelle 1 sind die Korrelationskoeffizienten zum ersten Messzeitpunkt der Testmotivation t1 dargestellt.

In der Experimentalgruppe bestanden zu diesem Zeitpunkt mäßige, aber signifikante Zusammenhänge zwischen den Ergebnissen im NV-MSVT und den FEW-2-Subtests Lage im Raum ( $r = 0,40^*$ ), Abzeichnen ( $r = 0,40^*$ ) und räumliche Beziehungen ( $r = 0,41^*$ ). Diese Korrelationen waren in der Kontrollgruppe nicht signifikant. Tabelle 2 zeigt die Korrelationskoeffizienten zum zweiten Messzeitpunkt der Testmotivation t2.

In der Experimentalgruppe bestand zu diesem Zeitpunkt für die Ergebnisse aller Subtests des FEW-2 ein signifikanter Zusammenhang zu den Ergebnissen des NV-MSVT. Die Subtests Auge-Hand-Koordination ( $r = 0,38^*$ ), Figur-Grund-Wahrnehmung ( $r = 0,45^*$ ), Gestaltschluss ( $r = 0,35^*$ ) und visuo-motorische Geschwindigkeit ( $r = 0,41^*$ ) zeigten einen mäßig signifikanten

Zusammenhang zum NV-MSVT-Gesamtscore; für die Subtests Lage im Raum ( $r = 0,65^{**}$ ), Abzeichnen ( $r = 0,48^{**}$ ), räumliche Beziehungswahrnehmung ( $r = 0,55^{**}$ ) und Formkonstanz ( $r = 0,46^{**}$ ) ergaben sich hochsignifikante Korrelationen. In der Kontrollgruppe bestand lediglich ein mäßig signifikanter Zusammenhang beim Subtest Figur-Grund-Wahrnehmung ( $r = 0,35^*$ ) und hochsignifikante Korrelationen bei den Subtests Abzeichnen ( $r = 0,45^*$ ), räumliche Beziehungswahrnehmung ( $r = 0,42^{**}$ ) und Formkonstanz ( $r = 0,52^{**}$ ).

Weiterhin wurde der Einfluss der Belohnung auf die Testmotivation durch einen Vergleich der Ergebnisse des FEW-2 bei der Experimental- und Kontrollgruppe mittels des t-Tests untersucht (siehe Tabelle 3 und 4).

Weder zum ersten Messzeitpunkt t1 noch zum zweiten Messzeitpunkt t2 ergaben sich signifikante Unterschiede zwischen den Mittelwerten der Experimental- und Kontrollgruppe.

Tabelle 2: Korrelationskoeffizienten zwischen den einzelnen FEW-2-Subtestergebnissen und dem Gesamtscore des NV-MSVT zum zweiten Messzeitpunkt t2.

Subtests FEW-2	Experimentalgruppe	Kontrollgruppe
	(mit Belohnung)	(ohne Belohnung)
Auge-Hand-Koordination	0,38*	-0,10
Lage im Raum	0,65**	0,25
Abzeichnen	0,48**	0,45**
Figur-Grund-Wahrnehmung	0,45*	0,35*
Räumliche Beziehungswahrnehmung	0,55**	0,42**
Gestaltschluss	0,35*	0,21
Visuo-motorische Geschwindigkeit	0,41*	-0,04
Formkonstanz	0,46**	0,52**

Legende: Signifikanz auf 5%-Niveau (\*), 1%-Niveau (\*\*)

Tabelle 3: Einfluss der Belohnung auf die Testergebnisse der NV-MSVT-Subtests zum ersten Messzeitpunkt der Testmotivation t1.

	Experimentalgruppe		Kontrollgruppe		t-Test	
	M	SD	M	SD	t-Wert	p-Wert
NV-MSVT-Gesamtscore	87,81	10,40	84,89	9,28	1,32	0,19
IR	98,86	2,99	98,56	4,84	0,32	0,75
DR	92,57	12,74	88,78	12,07	1,36	0,18
CON	91,71	14,19	88,67	11,75	1,05	0,30
DRA	70,57	17,77	60,22	19,22	1,04	0,30
DRV	96,29	6,46	95,78	8,39	0,30	0,77
PA	96,57	7,25	97,78	6,36	-0,79	0,43
FR	34,86	15,22	33,89	15,55	0,28	0,78

Legende: M = Mittelwert, SD = Standardabweichung, IR: Immediate Recognition Trial; DR: Delayed Recognition Trial; CON: Consistency; DRA: Delayed Recognition Trial/Archetypes; DRV: Delayed Recognition Trial/Variations; PA: Paired Associate Recall; FR: Free Recall

## Diskussion

### Beantwortung der Fragestellung

Bereits Green et al. stellten 2001 fest, dass ein Zusammenhang zwischen der Testmotivation und den Ergebnissen neuropsychologischer Leistungstests besteht. Da es sich beim FEW-2 ebenfalls um ein leistungsorientiertes neuropsychologisches Verfahren handelt und praktische Erfahrungen gezeigt haben, dass v. a. jüngere Kinder die Aufgaben scheinbar häufig mit mangelhafter Testmotivation bearbeiten, ergab sich die Frage, ob die Ergebnisse des FEW-2 durch die Motivation (gemessen über NV-MSVT) verändert werden. Die Messung der Testmotivation erfolgte zu zwei verschiedenen Zeitpunkten (vor t1 und nach t2), da davon auszugehen ist, dass sie sich während der

30 bis 45 Minuten dauernden Erhebungsphase ändert. Darüber hinaus sollte untersucht werden, ob möglicherweise die Motivation durch die Ankündigung einer Belohnung beeinflussbar ist. Für die Kinder, denen ein Anreiz in Aussicht gestellt wurde (Experimentalgruppe), und für die Kinder, die kein solches Versprechen erhielten (Kontrollgruppe), erfolgte die Datenauswertung getrennt. Die Korrelationsanalysen zwischen den Ergebnissen des FEW-2 und NV-MSVT zum Zeitpunkt t1 ergaben für die Experimentalgruppe insgesamt einen stärkeren positiven Zusammenhang als für die Kontrollgruppe. Kinder mit einer höheren Testmotivation (gute Ergebnisse im NV-MSVT) erzielten tendenziell auch bessere Ergebnisse in den einzelnen Subtests des FEW-2. Diese Korrelation war aber nur für die Subtests Lage im Raum ( $r = 0,40^*$ ),

Tabelle 4: Einfluss der Belohnung auf die Testergebnisse der NV-MSVT-Subtests zum zweiten Messzeitpunkt der Testmotivation t2.

	Experimentalgruppe		Kontrollgruppe		t-Test	
	M	SD	M	SD	t-Wert	p-Wert
NV-MSVT-Gesamtscore	86,49	9,87	83,72	12,02	1,10	0,28
IR	98,57	2,86	97,95	4,98	0,65	0,52
DR	93,43	11,43	90,91	13,69	0,87	0,39
CON	92,29	12,45	90,00	13,68	0,77	0,45
DRA	64,14	17,76	60,00	17,72	1,03	0,31
DRV	96,00	8,81	93,86	16,31	0,70	0,49
PA	98,29	7,07	98,18	6,57	0,07	0,95
FR	34,86	14,58	31,36	11,83	1,18	0,24

Legende: M = Mittelwert, SD = Standardabweichung, IR: Immediate Recognition Trial; DR: Delayed Recognition Trial; CON: Consistency; DRA: Delayed Recognition Trial/Archetypes; DRV: Delayed Recognition Trial/Variations; PA: Paired Associate Recall; FR: Free Recall

Abzeichnen ( $r = 0,40^*$ ) und räumliche Beziehungen ( $r = 0,41^*$ ) mäßig signifikant. In der Kontrollgruppe ließen sich keine signifikanten Zusammenhänge nachweisen.

Zum Zeitpunkt t2 wurde eine etwas deutlichere Korrelation zwischen den Ergebnissen im FEW-2 und denen im NV-MSVT in beiden Untersuchungsgruppen festgestellt, als dies zum Zeitpunkt t1 beobachtbar war. In der Experimentalgruppe waren die Ergebnisse für alle FEW-2-Subtests mäßig bzw. hochsignifikant (Auge-Hand-Koordination [ $r = 0,38^*$ ], Lage im Raum [ $r = 0,65^{**}$ ], Abzeichnen [ $r = 0,48^{**}$ ], Figur-Grund-Wahrnehmung [ $r = 0,45^*$ ], räumliche Beziehungswahrnehmung [ $r = 0,55^{**}$ ], Gestaltschluss [ $r = 0,35^*$ ], visuo-motorische Geschwindigkeit [ $r = 0,41^*$ ], Formkonstanz [ $r = 0,46^{**}$ ]). Auch in der Kontrollgruppe verstärkte sich die Korrelation

zwischen den Ergebnissen der beiden Tests. Im Subtest Figur-Grund-Wahrnehmung ( $r = 0,35^*$ ) ergab sich ein mäßig signifikanter Zusammenhang. In den Subtests Abzeichnen ( $r = 0,45^{**}$ ), räumliche Beziehungen ( $r = 0,42^{**}$ ) und Formkonstanz ( $r = 0,52^{**}$ ) wurden hochsignifikante Zusammenhänge ermittelt. Der mit diesen Ergebnissen nachgewiesene Einfluss der Testmotivation auf die Ergebnisse des FEW-2 unterstreicht die Dringlichkeit der eingangs erwähnten Empfehlung, Überprüfungen der Motivation routinemäßig in neuropsychologische Testungen zu integrieren, da aufgrund (verfälschter) diagnostischer Ergebnisse therapeutische Interventionen geplant und durchgeführt werden. Die Daten belegen weiter, dass der Zeitpunkt, zu dem die Motivationsmessung erfolgt, die Aussagekraft der Analyse stark beeinflusst.

Die Annahme, dass sich die Testmotivation und das Abschneiden im FEW-2 durch die Ankündigung einer Belohnung verbessern, konnte in dieser Studie nicht bestätigt werden. Sowohl für den Gesamtscore der NV-MSVT als auch für die einzelnen Subtests gab es zu beiden Motivationsmesszeitpunkten (t1 und t2) keine signifikanten Mittelwertdifferenzen zwischen den Kindern der Experimental- und der Kontrollgruppe. Daraus resultierend verbesserten sich die Testergebnisse im FEW-2 durch die Ankündigung der Belohnung ebenfalls nicht.

### **Methodenkritische Reflexion**

Im Rahmen der Belohnungsankündigung wurden die Mädchen und Jungen der Experimentalgruppe aufgefordert, sich so gut wie möglich anzustrengen. Für die Kinder der Kontrollgruppe gab es keine entsprechende Instruktion. Diese Aufforderung ist möglicherweise Mitursache für die unterschiedlichen Ergebnisse der Korrelationsanalyse in den beiden Gruppen. Leistungsbereitschaft spielt besonders in denjenigen FEW-2-Subtests eine entscheidende Rolle, die für die Kinder einen hohen kognitiven Anspruch aufweisen (Lage im Raum, Abzeichnen und räumliche Beziehungswahrnehmung). Ist den Mädchen und Jungen nicht bewusst, dass von ihnen eine möglichst gute Leistung erwartet wird (Kontrollgruppe), dann führt eine gute Testmotivation im Sinne von „Mitmachen“ nicht zwangsläufig zu besseren Testergebnissen im FEW-2. Diese fehlende Anstrengung der Kinder der Kontrollgruppe, ihre beste Leistung zu zeigen, erklärt möglicherweise die im Vergleich zur Experimentalgruppe schlechtere Korrelation zwischen den beiden Tests.

Ursächlich für die besseren Korrelationen zum zweiten Messzeitpunkt der Testmotivation (t2) könnte eine Abhängigkeit von der Konzentrationsleistung sein, da sich die Bedeutung der Tests für die Kinder mit abnehmender Konzentrationsleistung verrin-

gert. Die Kinder sehen keinen Nutzen mehr in ihrem Handeln (Baumert & Demmrich, 2001). Je konzentrierter ein Kind am Ende des FEW-2 ist, desto höher ist möglicherweise auch seine Testmotivation im Sinne von Leistungsbereitschaft, was als Folge auch die Ergebnisse im FEW-2 positiv beeinflusst. Daher wird es am Ende des FEW-2 für ein gutes Testergebnis zunehmend wichtiger, dass die Kinder noch gut motiviert sind, ihre beste Leistung zu zeigen. Zusammenfassend scheinen die Ergebnisse der vorliegenden Studie zu belegen, dass es bei der Durchführung des FEW-2 von Bedeutung ist, ob man die Testmotivation der Kinder im Sinne von „Mitmachen“ oder im Sinne von „Leistungsbereitschaft“ untersucht. Bei der Motivation zur Leistungsbereitschaft verbessern sich die Ergebnisse im FEW-2. Leistungsbereitschaft stellt somit eine Voraussetzung für valide Testergebnisse im FEW-2 dar. Diese Hypothese ist wenig überraschend, da sowohl der NV-MSVT als auch der FEW-2 Leistungstests darstellen, bei denen neuropsychologische Leistungen von den Kindern abverlangt werden.

Eventuelle Ursachen für das Ausbleiben des erwarteten positiven Effekts der Belohnung könnten das Auftreten von Anreiz- oder Wirksamkeitsdefiziten oder der hohe Anspruch an das Konzentrationsvermögen sein. Möglicherweise hatte die in Aussicht gestellte Belohnung nicht den intendierten Anreiz für die Kinder (Anreizdefizit). In diesem Fall entsprach die Belohnung nicht einem angemessenen Ergebnis, infolgedessen sich die Testmotivation der Kinder auch nicht erhöhen konnte (Rheinberg, 2004). Bestand zwar kein Anreizdefizit, aber ein Wirksamkeitsdefizit, könnte die Testmotivation der Kinder unter dem hohen Konzentrationsanspruch, der Komplexität der Aufgaben oder der Aufgabenanzahl im FEW-2 gelitten haben, wodurch die Belohnung als Motivationsförderung unwirksam wurde. Jedoch lässt sich über die Ursache für das Ausbleiben des Belohnungseffektes

nur spekulieren, da die Analyse nicht zwischen verschiedenen Motivationsformen bei den Kindern der Experimentalgruppe unterscheidet.

### **Praktische Implikationen**

Basierend auf den vorliegenden Untersuchungsergebnissen kann abschließend festgehalten werden, dass sich in der Experimentalgruppe in allen Subtests mäßig bis hochsignifikante Zusammenhänge zwischen der Testmotivation am Ende des FEW-2 (t2) und den Testergebnissen des FEW-2 ergaben. Das bevorzugte Auftreten von Korrelationen in der Experimentalgruppe weist darauf hin, dass es von Bedeutung ist, ob die Testmotivation der Kinder beim Bearbeiten des FEW-2 im Sinne von „Mitmachen“ oder im Sinne von „Leistungsbereitschaft“ angesprochen wird. Eine hohe Motivation im Sinne von Leistungsbereitschaft führt zu besseren Ergebnissen im FEW-2. Als Ursache für die bessere Korrelation zwischen dem FEW-2 und dem Motivationstest zum zweiten Messzeitpunkt der Testmotivation (t2) kommt eine Abhängigkeit der Testmotivation von der Konzentrationsleistung in Frage. Anreiz- oder Wirksamkeitsdefizite oder der lang andauernde Konzentrationsanspruch sind dagegen möglicherweise für den fehlenden positiven Effekt einer Belohnung verantwortlich.

Aus den vorliegenden Ergebnissen ergeben sich für die Praxis verschiedene Handlungsempfehlungen. Primär ist es wichtig, die Testmotivation im Sinne von Leistungsbereitschaft sicherzustellen. Um möglichen Wirksamkeitsdefiziten entgegenzuwirken, ist jedoch auch das Testmaterial zu überdenken. Damit sich der Konzentrationsanspruch des FEW-2 verringert, könnten zusätzliche Pausen eingelegt, der Test an zwei Tagen durchgeführt oder die Anzahl der Aufgaben in den einzelnen Untertests verringert werden. Um die Kinder nicht nur zum Mitmachen, sondern auch zur Leis-

tung zu motivieren, sind sie bei einer Belohnungsankündigung immer ausdrücklich dazu aufzufordern, die beste Leistung zu zeigen. Nur so sind valide Testergebnisse sicherzustellen. Grundsätzlich sollte sich jede Diagnostikerin bzw. jeder Diagnostiker immer an die Empfehlungen zur Durchführung von psychologischen Testverfahren bei Kindern von Petermann und Macha (2005) halten und die motivationsfördernden Maßnahmen von Rheinberg (2004) berücksichtigen, um eine optimale Testmotivation im Sinne von Leistungsbereitschaft zu unterstützen.

### **Ansatzpunkte für künftige Forschungsarbeiten**

Wie oben beschrieben, ist eine ausreichende Testmotivation im Sinne von Leistungsbereitschaft für valide Ergebnisse im FEW-2 unverzichtbar. In diesem Kontext wären Untersuchungen zur Art der angekündigten Belohnung wünschenswert. Möglicherweise ist es sinnvoll, den Kindern eine Belohnung in Aussicht zu stellen, ohne diese näher zu definieren. Dadurch könnte das Neugierverhalten der Kinder anregt und Anreizdefizite vermieden werden, falls die angekündigte Belohnung nicht den Wünschen des Kindes entspricht. Durch ein solches Vorgehen wäre es eventuell möglich, einem vollständigen Motivationsdefizit vorzubeugen.

### **Literatur**

- Allen, L.M., Conder, R.L., Green, P. & Cox, D.R. (1997). *Manual for the Computerized Assessment of Response Bias (CARB)*. Durham, NC: CogniSyst.
- Baumert, J. & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16, 441–462.
- Brown, A.L. (1975). *The development of memory: Knowing, knowing about knowing, and*

- knowing how to know. In H.W. Reese (Ed.), *Advances in child development and behaviour* (pp. 104–116). New York: Academic Press.
- Büttner, G., Dacheneder, W., Schneider, W. & Weyer, K. (2008). *FEW-2: Frostigs Entwicklungstest zur visuellen Wahrnehmung – 2*. Göttingen: Hogrefe.
- Constantinou, M., Bauer, L. & Ashendorf, L. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tasks? *Archives of Clinical Neuropsychology*, 20, 191–198.
- Constantinou, M. & McCaffrey, R.J. (2003). Using the TOMM for evaluating children's effort to perform optimally on neuropsychological measures. *Child Neuropsychology*, 9, 81–90.
- Courtney, J.C., Dinkins, J.P., Allen, L.M. & Kuroski, K. (2003). Age related effects in children taking the Computerized Assessment of response Bias and the Word Memory test. *Child Neuropsychology*, 9, 109–116.
- Green, P. (2004). Testmotivation und ihre Messung. *Report Psychologie*, 29, 303–308.
- Green, P. (2007). The Pervasive Influence of Effort on Neuropsychological Tests. *Physical Medicine and Rehabilitation Clinics of North America*, 18, 43–68.
- Green, P. (2008). *Green's Non-Verbal Medical Symptom Validity Test (NV-MSVT). User's Manual for Microsoft Windows*. Edmonton: Green's Publishing.
- Green, P., Allen, L. & Astner, K. (1996). *The Word Memory Test (WMT)*. Durham, NC: CogniSyst.
- Green, P. & Flaro, L. (2003). Word Memory Test Performance in children. *Child Neuropsychology*, 9, 189–207.
- Green, P., Gervais, R. & Flaro, L. (2006). Comparing the Non-Verbal MSVT with the TOMM, WMT, MSVT & RDS. *San Antonio Meeting of the National Academy of Neuropsychology*.
- Green, P., Rohling, M.L., Lees-Haley, P.R. & Allen, L.M. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, 15, 1045–1060.
- Heckhausen, H. (1991). *Motivation and action*. Berlin: Springer.
- Heise, D., Saupe, J. & Johnenning, A. (2008). *Die Selbstbestimmungstheorie der Motivation*. Unveröffentlichtes Manuskript, Friedrich-Schiller-Universität Jena.
- Iverson, G.L. (2006). Ethical Issues Associated With the Assessment of Exaggeration, Poor Effort, and Malingering. *Applied Neuropsychology*, 13, 77–90.
- Kuhl, J. (1996). *Wille und Freiheitserleben. Formen der Selbststeuerung*. In J. Kuhl & H. Heckhausen (Hrsg.), *Motivation, Volition und Handlung* (S. 665–768). Göttingen: Hogrefe.
- Merten, T. (2002). Fragen der neuropsychologischen Diagnostik bei Simulationsverdacht. *Fortschritte der Neurologie-Psychiatrie*, 70, 126–138.
- Merten, T. & Brockhaus, R. (2004). Ein Paradigmenwechsel in der Neuropsychologie. *Report Psychologie*, 29, 302–308.
- Perlmutter, M. & Lange, G.A. (1978). *A developmental analysis of recall-recognition distinctions*. In P.A. Ornstein (Ed.), *Memory development in children* (pp. 243–258). Hilldale, NJ: Erlbaum.
- Petermann, F. & Macha, T. (2005). *Psychologische Tests für Kinderärzte*. Göttingen: Hogrefe.
- Piaget, J. & Inhelder, B. (1973). *Memory and Intelligence*. New York: Basic Books.
- Reitan, R.M. & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.
- Rey A. (1958). *L'examen Clinique en Psychologie*. Paris: Presses Universitaires de France.
- Rheinberg, F. (1995). *Motivation*. Stuttgart: Kohlhammer.
- Rheinberg, F. (2004). *Motivationsdiagnostik. Kompendien. Psychologische Diagnostik*. Göttingen: Hogrefe.
- Rohling, M.L. (2004). Who do They Think They're Kidding: A Review of the Use of Symptom Validity Tests with Children. *Division of Clinical Neuropsychology, Newsletter 40, American Psychological Association Newsletter*, 22, 1–8.
- Sokolowski, K. (1993). *Emotion und Volition*. Göttingen: Hogrefe.
- Sophian, C. & Stigler, J.W. (1981). Does Recognition Memory Improve with Age? *Journal of Experimental Child Psychology*, 32, 343–353.
- Tombaugh, T.N. (1996). *Test of Memory Malinger (TOMM)*. San Antonio, TX: Pearson.



- Wang, M.C., Haertel, G.D. & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249–294.
- Weiner, B. (1994). *Motivationspsychologie*. Weinheim: Psychologie Verlags Union.
- Wigfield, A. & Eccles, J.S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, 12, 265–310.

## **Anschriften der Autoren**

PEGGY BEUTHAN  
Auweg 9a  
36456 Barchfeld  
peggy\_menz@web.de

ASS.-PROF. DR. WILLI GESER  
Leopold-Franzens-Universität Innsbruck  
Institute for Psychology  
Innrain 52  
A-6020 Innsbruck, Austria  
willi.geser@uibk.ac.at

UNIV.-DOZ. DR. CLAUDIA SCHUSTER-  
SCHITZ  
University for Health Sciences, Medical  
Informatics and Technology  
Department for Human and Economic  
Sciences  
Eduard-Wallnöfer-Zentrum 1  
A-6060 Hall in Tirol, Austria  
claudia.schusterschitz@umit.at