



 **Universität Trier**

*Fachbereich I - Psychologie
Zentrum für psychologische Diagnostik, Begutachtung und Evaluation (ZDiag)*

Fördert Schachunterricht in der Grundschule die geistige Entwicklung der Kinder?

Abschlussbericht über eine Evaluationsstudie zum
Schachunterricht in einer Trierer Grundschule

mit Unterstützung der Deutschen Schulschachstiftung, der
Deutschen Schachjugend und der Nikolaus Koch-Stiftung Trier

Prof. Dr. Sigrun-Heide Filipp

Dipl.-Psych. Holger Spieles

Februar 2007

Dr. Thomas Wehr sowie den (damaligen) studentischen Mitarbeiterinnen und Mitarbeitern Anne Krämer, Johannes Kregel, Christian Meisenzahl, Jutta Propson, Achim Schneider und Janne Strakeljahn sei für wertvolle Unterstützung gedankt.

Zusammenfassung

Im Rahmen der Qualitätsentwicklung an Schulen in Rheinland-Pfalz hat die Grundschule Trier-Olewig in Zusammenarbeit mit der „Deutschen Schulschachstiftung“ seit 2003 Schachunterricht in den wöchentlichen Stundenplan aller vier Klassenstufen integriert. Das „Zentrum für Psychologische Diagnostik, Begutachtung und Evaluation“ (ZDiag) erhielt den Auftrag zu prüfen, inwieweit der Schachunterricht positive Auswirkungen auf die (insbesondere kognitive) Entwicklung der Schülerinnen und Schüler der ersten vier Klassenstufen zeitigt. Um einen solchen Nachweis prinzipiell führen zu können, wurde neben dieser „Experimental-Schule“ mit Schachunterricht eine „Kontroll-Schule“ ohne Schachunterricht, die sich in anderen relevanten Merkmalen (z.B. sozioökonomischer Hintergrund der Kinder, Klassengrößen) nicht von der „Experimentalschule“ unterscheiden sollte, in die Studie einbezogen. Durch den angestrebten Vergleich der beiden Schulen sollte es möglich sein, die bei Kindern dieses Alters allgemein zu erwartenden Entwicklungsfortschritte von den spezifischen Entwicklungsfortschritten zu trennen, die sich auf den Schachunterricht in der Experimental-Schule zurückführen lassen. Die Studie war so angelegt, dass die Kinder des ersten, zweiten, dritten und vierten Schuljahres in beiden Schulen mittels psychologischer Testverfahren untersucht wurden, und zwar beginnend vom Oktober 2003 bis zum Juli 2006 zu insgesamt fünf Testzeitpunkten. Mit den Testverfahren sollten ausgewählte Indikatoren der intellektuellen Leistungsfähigkeit und der Konzentrationsfähigkeit, schulische Leistungsfähigkeit in zwei Bereichen sowie Aspekte der seitens der Kinder erlebten Integration in die Schule erfasst werden. Die Ergebnisse der Evaluationsstudie verweisen auf ein komplexes Befundbild, das sich kaum in wenigen Sätzen zusammenfassen lässt. Die einzelnen Befunde differieren deutlich in Abhängigkeit davon, welchen Leistungsaspekt, welche Klassenstufe und welchen Testzeitpunkt man betrachtet. Einige der Befunde deuten darauf hin, dass der Schulschachunterricht (insbesondere mit Blick auf die beiden jüngeren Klassenstufen) förderliche Effekte auf die Entwicklung der Kinder gezeitigt hat, während eine Reihe anderer Einzelbefunde dafür spricht, dass es zwischen den Kindern beider Schulen (mit und ohne Schachunterricht) keine systematischen Unterschiede im Niveau und im Verlauf ihrer Leistungsfähigkeit gibt. In einem abschließenden Kommentar werden Einschränkungen der Evaluationsstudie thematisiert.

INHALTSÜBERSICHT

1	Stand der Forschung	4
2	Problemaufriss und Ziele der Evaluationsstudie	9
3	Ausgangssituation und Rekrutierung der Schülerstichprobe	11
4	Zeitplan der Erhebungen	12
5	Stichprobenbeschreibung	14
6	Testbeschreibungen	17
7	Auswertungshinweise	28
8	Ergebnisse	29
9	Zusammenfassende Betrachtung des Befundmaterials und abschließender Kommentar	104
	Anhang	109

1 Stand der Forschung

Schachspielen wird gemeinhin mit Ausdauer und Geduld, hoher Konzentration, strategischem Denken, motorischer Ruhe und vielen anderen positiv bewerteten Verhaltensaspekten in Verbindung gebracht. Wenn Kinder frühzeitig das Schachspiel erlernen, dann sollte sich dies – so lautet die durchaus plausible Erwartung – im Aufbau entsprechender Verhaltensmerkmale niederschlagen; Schachspiel sollte aber auch die kognitive Leistungsfähigkeit von Kindern steigern, und es sollten daraus wichtige Impulse für die kindliche Entwicklung insgesamt erwachsen. Diese Überlegungen sind zweifelsohne Leitgedanken der „Deutschen Schulschachstiftung e. V.“, die es sich zum Ziel gesetzt hat, die Tradition des Schulschachs auszubauen und zu pflegen.¹ Sie dürften auch den durch Gerlinde und Hans-Georg Mehlhorn gegründeten „Kreativschulen“ zugrunde liegen, in denen das Schachspiel als regelmäßigen Bestandteil in den Schulunterricht eingeführt wurde.² So naheliegend all diese Überlegungen auch immer sein mögen, so dürftig ist unser diesbezügliches Wissen aus *systematischen* Studien.

Eine aktuelle Literaturrecherche in den Datenbanken der nationalen und internationalen psychologischen Fachliteratur (PSYINDEX und PsycInfo) bestätigt, dass sich die psychologische Forschung bislang nur sehr wenig mit den Auswirkungen des Schachspiels auf die kognitive Entwicklung von Kindern beschäftigt hat. Eine Kombination der Suchbegriffe „chess“ „children“ und „development“ ergab in der international gültigen Datenbank PsycInfo insgesamt 204 Treffer. Eine entsprechende Literaturrecherche zu Beginn des Jahres 2005 hatte 165 Treffer ergeben. Die Suche nach einer Kombination der Schlüsselbegriffe „children“, „chess“ und „cognitive development“ reduzierte die Trefferzahl nochmals auf 12 (2005: 10 Treffer). Bei genauerer Inspektion der Trefferlisten erwies sich keine der enthaltenen Arbeiten als i.e.S. relevant für die in der vorliegenden Evaluationsstudie betrachtete Thematik. Zudem mussten drei Studien aus Russland bzw. Ungarn, die in der jeweiligen Landessprache verfasst sind, ausgeschlossen werden.

Bei den im weiteren Sinne als relevant eingestuften Studien handelt es sich zum einen um eine Arbeit von Schneider, Gruber, Gold und Opwis aus dem Jahre 1993,

¹ Siehe <http://www.schulschachstiftung.de>

² siehe <http://www.spiegel.de/schulspiegel/wissen/0,1518,445883,00.html>

die unter dem Titel „*Chess expertise and memory for chess positions in children and adults*“ veröffentlicht wurde. Daneben lassen sich auch von der Studie von Gruber, Renkl und Schneider (1994) mit dem Titel „*Expertise und Gedächtnisentwicklung: Längsschnittliche Befunde aus der Domäne Schach*“ Bezüge zu der vorliegenden Evaluationsstudie herstellen³. Die wesentlichen Ergebnisse beider Studien sollen nachfolgend kurz dargestellt werden.

Die Studie von Schneider et al. (1993) stellt eine Replikation und Erweiterung einer klassischen Studie von Chi (1978) zur Expertise im Schach dar. Entwicklungspsychologisch ist im Allgemeinen davon auszugehen, dass sich die Leistungen im kognitiven Bereich mit zunehmendem Alter verbessern und dass Erwachsene in ihrer kognitiven Leistungsfähigkeit Kindern überlegen sind. Das gilt insbesondere auch für Gedächtnisleistungen. Als ein wesentliches Ergebnis der Forschung von Chi (1978) kann deshalb die Erkenntnis gelten, dass sich diese Überlegenheit umkehren lässt und Kinder durchaus bessere Leistungen erzielen können als Erwachsene. Im vorliegenden Falle handelt es sich nämlich um Kinder, die "Experten" im Schachspiel waren und die deutlich bessere Gedächtnisleistungen für Schachkonstellationen zeigten als Erwachsene, die nicht Schach spielten (und hier als „Novizen“ bezeichnet wurden). Das Hauptziel der Studie von Schneider et al. (1993) war es, darüber hinaus zu prüfen, wie bestimmte Merkmale der Gedächtnisaufgaben die Erinnerungsleistungen der Probanden beeinflussen. Dabei handelte es sich um die Vertrautheit mit der Konstellation der Schachfiguren (sinnvolle vs. zufällige Konstellationen) und um die Vertrautheit sowohl mit der geometrischen Struktur des Schachbrettes sowie der Form und Farbe der Schachfiguren (siehe unten). In die Studie einbezogen waren 40 Erwachsene (Altersdurchschnitt: 26,8 Jahre) und 40 Kinder (Altersdurchschnitt: 11,9 Jahre). Diese beiden Altersgruppen setzten sich je zur Hälfte aus „Experten“ und „Novizen“ im Schachspiel zusammen, so dass die folgenden vier Gruppen mit jeweils 20 Probanden resultierten: Experten im Kindesalter, Novizen im Kindesalter,

³ Die vollständigen Literaturangaben der oben genannten Studien lauten:

- (1) Schneider, W.; Gruber, H.; Gold, A. & Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. *Journal of Experimental Child Psychology*, 56(3), 328-349.
- (2) Gruber, H.; Renkl, A. & Schneider, W. (1994). Expertise und Gedächtnisentwicklung. Längsschnittliche Befunde aus der Domäne Schach. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 26(1), 53-70.

erwachsene Experten und erwachsene Novizen. Gemessen wurden Gedächtnisleistungen für sinnvolle und zufällige Schachkonstellationen, wobei jeweils das unmittelbare Behalten der vorgegebenen Muster nach der ersten Präsentation erfasst wurde sowie die Verbesserung der Gedächtnisleistungen über mehrere Durchgänge hinweg (d.h. den Probanden wurde jedes Muster vier weitere Male präsentiert, um die Verbesserung ihrer Gedächtnisleistungen zu registrieren). Als weitere Gedächtnisaufgabe kam ein erfundenes „Brettspiel“ hinzu, das darin bestand, dass die Probanden nun die Positionen von Holzstückchen (die in Form und Farbe sich klar von Schachfiguren unterschieden) erinnern mussten, wobei diese zusätzlich Holzstückchen auf einem Spielbrett angeordnet waren, das verschiedene geometrische Formen (Kreise, Dreiecke, etc.) und nicht die üblichen quadratischen Felder eines Schachbretts aufwies. Auch bei dieser ersten „Kontrollaufgabe“ wurden das unmittelbare Behalten und die Verbesserung der Gedächtnisleistung über mehrere Durchgänge hinweg registriert. Eine weitere Kontrollaufgabe bestand darin, dass die Probanden Zahlenreihen erinnern sollten.

Die wichtigste Annahme der Autoren war nun, dass die Überlegenheit der Experten (egal welchen Alters) bei den sinnvollen Schachkonstellationen am deutlichsten ausfallen und bei den zufälligen Konstellationen zwar geringer ausfallen sollte, aber immer noch vorhanden sein. Hinsichtlich der ersten Kontrollaufgabe wurde erwartet, dass es keine systematischen (d.h. überzufälligen) Unterschiede in den Gedächtnisleistungen der Experten und der Novizen geben sollte. Bei der zweiten Kontrollaufgabe sollten nur altersgebundene Unterschiede auftreten, d.h. dass hier die erwachsenen Probanden aufgrund ihrer höheren Gedächtnisspanne bessere Leistungen zeigen sollten als die Probanden im Kindesalter.

Die Autoren konnten zeigen, dass die Experten im Kindesalter bei der Rekonstruktion von Schachkonstellationen bessere Resultate erzielten als Novizen im Erwachsenenalter, und zwar für die sinnvollen *und* für die zufälligen Figurenanordnungen. Die Erwartung der Autoren, dass Schach spielende Kinder und Erwachsene (d.h. die Experten) bei Vorgabe des neu erfundenen Brettspiels mit den Holzstückchen (Kontrollaufgabe 1) keinen Vorteil mehr gegenüber den Schachnovizen besitzen sollten, konnte nur für das unmittelbare Behalten bestätigt werden. Bei der Betrachtung des Lernfortschritts über mehrere Durchgänge hinweg war wiederum ein positiver Einfluss der Expertise zu verzeichnen. Die Autoren

schlussfolgerten daraus, dass die Expertise in der Domäne Schach auch bei ähnlichen Aufgabenanforderungen hilfreich ist und das Erlernen neuer Anforderungen mit ähnlichem Charakter erleichtert. Dies wurde darauf zurückgeführt, dass Experten die Informationen aus der kurzen Lernphase tiefer verarbeiten und sie somit auch nach längerer Zeit vergleichsweise gut abrufen können. Fazit der Untersuchung war, dass sich Vorwissen (=Expertise) positiv auf Gedächtnisleistungen auswirkte, wenn das zu erinnernde Reizmaterial mit dem Bereich, in dem das entsprechende Vorwissen erworben worden war (hier: Schachspiel) in Zusammenhang stand. Kinder können Erwachsenen also im Hinblick auf die Gedächtnisleistungen überlegen sein, wenn sie ein größeres Wissen in dem betreffenden Bereich besitzen.

In der zweiten hier darzustellenden Studie (Gruber et al., 1994) wurden die Jugendlichen, die an der Studie von Schneider et al. (1993) teilgenommen hatten ein zweites Mal untersucht. Zum ersten Messzeitpunkt waren die Jugendlichen 12 Jahre und zum zweiten Messzeitpunkt 16 Jahre alt. Es ging zum einen um die Frage, inwieweit sich die Jugendlichen, die über den Untersuchungszeitraum mit dem Schachspiel aufgehört hatten, sich von den anderen Jugendlichen unterscheiden, die beim Schach spielen geblieben sind. Dazu waren die Jugendlichen bei der Erstmessung aufgrund von Indikatoren (z.B. Häufigkeit des Schachspielens, Zugehörigkeit zu einem Schachverein) als „Experten“ oder als „Novizen“ klassifiziert worden. Zum zweiten Messzeitpunkt wurde auf der Grundlage derselben Kriterien erneut eine Einstufung vorgenommen, um zu prüfen, ob sich der Expertisegrad zwischenzeitlich verändert hat. Diejenigen Jugendlichen, die zum ersten Messzeitpunkt als Experten, zum zweiten Messzeitpunkt jedoch als Novizen eingestuft wurden, bildeten die Gruppe der „Aussteiger aus der Expertisekarriere“. Das hier relevante Maß der Gedächtnisleistung entsprach den Variablen, wie sie schon in der Studie von Schneider et al. (1993) verwendet worden waren. Die „Aussteiger“ hatten bereits bei der Erstmessung schlechtere schachspezifische Gedächtnisleistungen gezeigt als die Jugendlichen, die beim Schachspiel geblieben waren und auch noch zum zweiten Erhebungszeitpunkt als Experten gelten konnten. Offensichtlich steigen gerade jene Jugendlichen aus der Expertisekarriere aus, die (verglichen mit den „Nicht-Aussteigern“) eher geringe Kompetenzen im Schachspiel besessen hatten, während bereits zum ersten Messzeitpunkt leistungsstarke Schachspieler das Interesse am Schachspiel nicht verlieren. Die zweite wesentliche

Frage in dieser Untersuchung war, welcher Zuwachs sich in den schachspezifischen und in den allgemeinen (schachunspezifischen) Gedächtnisleistungen über den betrachteten Zeitraum bei den zu t_2 als Experten bzw. als Novizen klassifizierten Kindern feststellen lässt.

Für die Experten zeigte sich ein Anstieg der Gedächtnisleistung von der Erst- zur Zweitmessung nur bei den schachspezifischen Gedächtnisaufgaben, nicht bei den Kontrollaufgaben. Die Autoren gehen davon aus, dass die Erfahrung der Experten mit dem zu erinnernden Material verantwortlich sein dürfte. Bei den Novizen zeigte sich ebenfalls ein Anstieg der schachspezifischen Gedächtnisleistung zwischen den beiden Messzeitpunkten. Darüber hinaus zeigten nur diese Kinder auch einen Zuwachs in der Gedächtnisspanne für Zahlen. Dieser Entwicklung soll den Autoren zufolge ein Anstieg der allgemeinen kognitiven Fähigkeiten der Probanden zugrunde liegen.

Die beiden Studien thematisieren - wie soeben dargelegt - vor allem die Bedeutung von Expertise in Schach für die Gedächtnisentwicklung und dokumentieren die entsprechende Überlegenheit in Gedächtnisleistungen („Experten-Novizen-Vergleich“). In diesem Zusammenhang ist es wichtig anzumerken, dass es im Rahmen der Expertiseforschung nicht primär darum geht, die Besonderheiten des Schachspiels selbst zu ermitteln, sondern eher um die Frage, wie sich hohes Vorwissen in einem Bereich auf diesbezügliche Gedächtnisleistungen auswirkt. In anderen Worten: Man könnte diese Forschung also mit Blick auf fast jede beliebige Wissensdomäne anwenden. Indes ist das Schachspiel ein geeignetes Mittel, um die Rolle des Vorwissens in einem Experiment prüfen zu können. Es bietet sich vor allem deshalb an, weil es Kinder gibt, die bereits „Experten“ sind, und es umgekehrt Erwachsene gibt, die als „Novizen“ zu betrachten sind. Damit ist es gerade in der Domäne Schach sehr gut möglich, eine Gruppe erwachsener Novizen mit einer Gruppe kindlicher Experten zu vergleichen. Inwieweit das Schachspiel bei Kindern (und Erwachsenen) eine Förderung der kognitiven Entwicklung allgemein, der Konzentrationsfähigkeit und Ausdauer bewirkt sowie ggf. auch Aspekte der subjektiv erlebten Integration in die Schule bewirkt, ist in der uns zugänglichen Fachliteratur nicht behandelt worden.

2 Problemaufriss und Ziele der Evaluationsstudie

In der vorliegenden Untersuchung sollte die Wirkung des Schachspiels vor allem auf die geistige Entwicklung von Schülerinnen und Schülern untersucht werden. Im Unterschied zu den beschriebenen Studien aus dem Kontext der Expertiseforschung waren die teilnehmenden Kinder im vorliegenden Fall keine Experten, sondern sie erlernten das Spiel im Rahmen des Schulunterrichts neu. Ein weiterer Unterschied zu den bisherigen Studien bestand darin, dass in der vorliegenden Arbeit das Hauptaugenmerk darauf lag, inwieweit das Schachspiel positive Auswirkungen auf die Entwicklung von Kindern hat. Insbesondere die folgenden Fragestellungen waren dabei von Interesse:

Hat das Erlernen des Schachspiels Auswirkungen auf basale kognitive Funktionen (v. a. Intelligenz, Konzentrationsfähigkeit) der Grundschüler? Können die durch das Schachspiel geförderten Fähigkeiten und Fertigkeiten auf allgemeine Anforderungen übertragen werden, d.h. lässt sich beispielsweise eine Förderung der Wahrnehmungsleistung und des schlussfolgernden Denkens nachweisen, welche sich dann in den Ergebnissen der jeweiligen Testverfahren niederschlagen sollte? Zeigen sich darüber hinaus eventuell auch Auswirkungen auf schulische Leistungen, die weniger eng mit den im Rahmen des Schachspiels bestehenden Anforderungen zusammenhängen, wie z.B. die Rechtschreibung oder auch die Integration der Schach spielenden Kinder in ihre schulische Umwelt? Wie verändert sich der (mögliche) Einfluss des Schachunterrichts über einen längeren Zeitraum hinweg? Aus diesen Fragestellungen wurden in enger Abstimmung mit der Leitung der Grundschule Olewig und anhand des Lernzielkatalogs der Schule die Ziele der Evaluation abgeleitet.

Unter Bezugnahme auf theoretische Überlegungen sollten sich förderliche Auswirkungen des Schachunterrichts auf folgende Entwicklungsmerkmale der Kinder nachweisen lassen:

- **Konzentrationsfähigkeit:** Unter Konzentrationsfähigkeit wird hier die Fähigkeit verstanden, sich über einen längeren Zeitraum effektiv mit einer Aufmerksamkeit erfordernden Aufgabe auseinander setzen zu können.

-
- Allgemeine intellektuelle Fähigkeiten: Zu erwarten war, dass hier vor allem die Aspekte des vernetzten Denkens und der räumlichen Wahrnehmungsfähigkeit durch die Erfahrungen im Schachspiel gefördert werden können. Konkret erfasst wurde die Fähigkeit der Kinder, in neuartigen Situationen und anhand von sprachfreiem, figuralem Material, Denkprobleme zu lösen, Beziehungen herzustellen, Regeln zu erkennen sowie relevante visuell dargebotene Merkmale zu identifizieren und rasch wahrzunehmen.
 - Rechtschreibfähigkeit: Hier war zu erwarten, dass Schachunterricht die Wahrnehmungsfähigkeit fördert und diese ggf. sich positiv auf Lesefähigkeiten und Leseverständnis und damit mittelbar auch auf die Rechtschreibung auswirken könnte⁴.

Darüber hinaus sollte geprüft werden, ob jenseits dieser leistungsbezogenen Aspekte Veränderungen in weiteren Merkmalen nachweisbar sind. Da der Lernzielkatalog der Schule unter anderem auch Leistungsmotivation und Sozialkompetenz thematisiert, wurde geprüft, inwieweit sich die zusätzlichen (sozialen) Anforderungen im Rahmen des Schachunterrichts auch auf ausgewählte Aspekte der schulischen Integration förderlich auswirken.

⁴ Eine entsprechende uns allerdings nicht zugängliche Studie, die in den USA durchgeführt worden sein soll und mehrfach in den Medien erwähnt wurde, hatte solche Zusammenhänge vorgeblich ermittelt.

3 Ausgangssituation und Rekrutierung der Schülerstichprobe

An der Untersuchung nahmen Schülerinnen und Schüler aus zwei Trierer Grundschulen teil. Die Grundschule Olewig (Experimentalschule) hatte eine Stunde Schachunterricht pro Woche in den Stundenplan aller vier Klassenstufen integriert, d.h. die Schülerinnen und Schüler wurden pro Woche in einer Schulstunde von ihren Lehrerinnen im Schachspiel unterwiesen. Die Lehrerinnen waren zuvor von Vertretern der „Deutschen Schulschachstiftung“ geschult worden, um den Kindern das Schachspiel kompetent vermitteln zu können. Dieses „Schulschach-Projekt“ wurde im Rahmen der Qualitätsentwicklung an Schulen in Rheinland-Pfalz ermöglicht und in Zusammenarbeit mit der „Deutschen Schulschachstiftung“ realisiert. Um die vermuteten Effekte des Schachspiels wissenschaftlich überprüfen zu lassen, beauftragte die Schulleitung der Grundschule Olewig das „Zentrum für Psychologische Diagnostik, Begutachtung und Evaluation (ZDiag)“ der Universität Trier mit der Durchführung der Evaluationsstudie. Finanziell gefördert wurde die Studie von der „Nikolaus-Koch-Stiftung“ sowie von der „Deutschen Schulschachstiftung“.

Um Aussagen über mögliche Wirkungen des Schachunterrichts treffen zu können, musste die Experimentalschule mit einer Kontrollschule verglichen werden; hierzu wurde eine Schule, die keinen Schachunterricht anbot, als Vergleichsschule gewählt. Durch diesen Vergleich sollte es möglich sein, die allgemein zu erwartenden Entwicklungsfortschritte im Grundschulalter von den Leistungsfortschritten, die durch den zusätzlichen Schulschachunterricht bewirkt sein könnten, zu trennen. Als Kontrollschule konnte die Grundschule Egbert gewonnen werden. Um zu überprüfen, ob etwaige positive Effekte des Schulschachunterrichts persistieren, d.h. sich über einen längeren Zeitraum auf die ausgewählten Aspekte der kognitiven Entwicklung auswirken, wurde die Untersuchung zudem als Längsschnittstudie über ein Drei-Jahres-Intervall mit mehrfachen Wiederholungsmessungen konzipiert (siehe Abschnitt 4).

Die Eltern der Schülerinnen und Schüler sowie das Lehrpersonal der beiden Schulen waren von Vertretern der Universität Trier über das Projekt und das zugrundeliegende Forschungsinteresse informiert worden. In die Studie wurden nur Schülerinnen und Schüler einbezogen, deren Eltern sich mit der Teilnahme ihrer

Kinder an der Untersuchung einverstanden erklärt hatten. Die Anonymität der Datengewinnung war gewährleistet, da alle verwendeten Arbeitsmaterialien lediglich mit den Initialen der Probanden kenntlich gemacht wurden. Die vollständigen Namen der teilnehmenden Schülerinnen und Schüler sind der Forschungsgruppe nicht bekannt.

4 Untersuchungsverfahren und Zeitplan der Erhebungen

In der Evaluationsstudie wurden im einzelnen folgende Testverfahren eingesetzt:

- *Differentieller Leistungstest–KG/KE (DL–KG/KE) zur Erfassung der Konzentrationsfähigkeit* von Kleber, Kleber & Hans (1999)
- *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* von Kurth & Büttner (1999)
- *Grundintelligenztest Skala 1 („Culture Fair Intelligence Test“; CFT 1)* von Weiß & Osterland (1997) resp. *Grundintelligenztest Skala 20 (CFT 20)* von Weiß (1998)
- *Hamburger Schreib-Probe für 1. bis 9. Klasse (HSP 1-9)* von May (2000)
- *Fragebogen zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6)* von Haeberlin et al. (1989)

In beiden Schulen wurden die Kinder zu insgesamt fünf Erhebungszeitpunkten untersucht. Die Erhebungen in den beiden Schulen fanden innerhalb der gleichen Woche statt, jeweils in zwei Schulstunden. Der erste Erhebungszeitpunkt lag zu Beginn des Schuljahres 2003/2004 (Anfang Oktober 2003), der zweite Erhebungszeitpunkt in der Mitte desselben Schuljahres, d.h. Anfang Februar 2004. Der dritte Erhebungszeitpunkt lag Ende Juni 2004, also am Ende des Schuljahres 2003/2004. Am Ende des Schuljahres 2004/2005 erfolgte die vierte Erhebung (Anfang Juli 2005), und ein Jahr später, d. h. Anfang Juli 2006, fand die fünfte und letzte Erhebung statt (siehe Abb. 5.1).

Zu *allen* Erhebungszeitpunkten wurden der „*Differentielle Leistungstest zur Erfassung der Leistungsfähigkeit*“ (DL-KG und DL/EG) sowie der „*Grundintelligenztest*“ (CFT 1 und CFT 20) jeweils in der der Altersstufe

angemessenen Version durch Mitarbeiter des ZDiag vorgegeben. Andere Testverfahren wurden nicht durchgängig eingesetzt:

Die Durchführung der „Hamburger Schreib-Probe“ (HSP) sowie der „Testreihe zur Prüfung der Konzentrationsfähigkeit“ (TPK) und des „Fragebogen zur Erfassung von Dimensionen der Integration von Schülern“ (FDI) unterschied sich in einigen Punkten von den vorgenannten Testverfahren. Erstere wurde nur zu vier Erhebungszeitpunkten, unabhängig von den beiden oben genannten Testverfahren (CFT; DL-KG resp. DL-EG) im Januar 2004, im Juni 2004, im Juli 2005 und im Juli 2006 vorgegeben. (Die Erhebung zum ersten Erhebungszeitpunkt im ersten Schuljahr war zeitverzögert durchgeführt worden, da anfangs noch keine für eine Testung hinreichenden Rechtschreibfähigkeiten vorausgesetzt werden konnten.) Die beiden anderen Verfahren (TPK und FDI) wurden insgesamt drei Mal erhoben: Im Juni 2004, Juli 2005 und Juli 2006. Die Durchführung dieser Verfahren wurde von den jeweiligen Klassenlehrerinnen übernommen, die die Testhefte und Fragebogen mit den entsprechenden Handanweisungen zur Durchführung erhalten hatten.

5 Stichprobenbeschreibung

Die Stichprobe bestand aus insgesamt 167 Schülern und Schülerinnen der Klassenstufen (1) bis (4) – im folgenden „Gruppe 1“, „Gruppe 2“, Gruppe 3“ und „Gruppe 4“ genannt. 84 Kinder besuchten die Grundschule in Olewig (Experimentalschule), 83 Kinder besuchten die Grundschule in Egbert (Kontrollschule); das mittlere Alter dieser Kinder aus den vier Klassenstufen zu den einzelnen Erhebungszeitpunkten ist der folgenden Tabelle 1 zu entnehmen.

Tabelle 1. Mittleres Alter der Schülerinnen und Schüler in beiden Schulen für die vier Gruppen und alle fünf Messzeitpunkte

Messzeitpunkt	Gruppe 1 (1. Schuljahr zu Studienbeginn)		Gruppe 2 (2. Schuljahr zu Studienbeginn)		Gruppe 3 (3. Schuljahr zu Studienbeginn)		Gruppe 4 (4. Schuljahr zu Studienbeginn)	
	Olewig n = 19	Egbert n = 22	Olewig n = 23	Egbert n = 22	Olewig n = 18	Egbert n = 20	Olewig n = 24	Egbert n = 19
	Oktober 2003 t1	6;6	6;6	7;4	7;7	8;7	8;5	9;8
Februar 2004 t2	7;0	7;0	7;8	8;0	9;1	8;9	10;2	10;0
Juni 2004 t3	7;4	7;3	8;2	8;4	9;5	9;3	10;5	10;4
Juli 2005 t4	8;4	8;3	9;2	9;4	10;5	10;3	-	-
Juli 2006 t5	9;4	9;3	10;2	10;4	-	-	-	-

Abbildung 5.1 veranschaulicht das sequentielle Untersuchungsdesign, gibt jeweils das mittlere Alter an und stellt dar, welche Klassenstufe die einzelnen Gruppen zu den jeweiligen Erhebungszeitpunkten besuchten. Durch den zwischenzeitlichen Wechsel auf weiterführende Schulen konnten jene Schülerinnen und Schüler, die zum ersten Erhebungszeitpunkt das dritte oder das vierte Schuljahr besucht hatten, zu den letzten Erhebungszeitpunkten wegen des zwischenzeitlich erfolgten Schulwechsels nicht mehr erreicht werden. Im einzelnen bedeutet dies, dass die Kinder, die zu Beginn der Untersuchung in der vierten Klassen war, weder im Juli 2005 noch im Juli 2006 nochmals getestet werden konnten; entsprechend waren die Kinder, die zum Beginn der Studie im dritten Schuljahr waren, für die letzte Datenerhebung im Juli 2006 nicht mehr erreichbar.

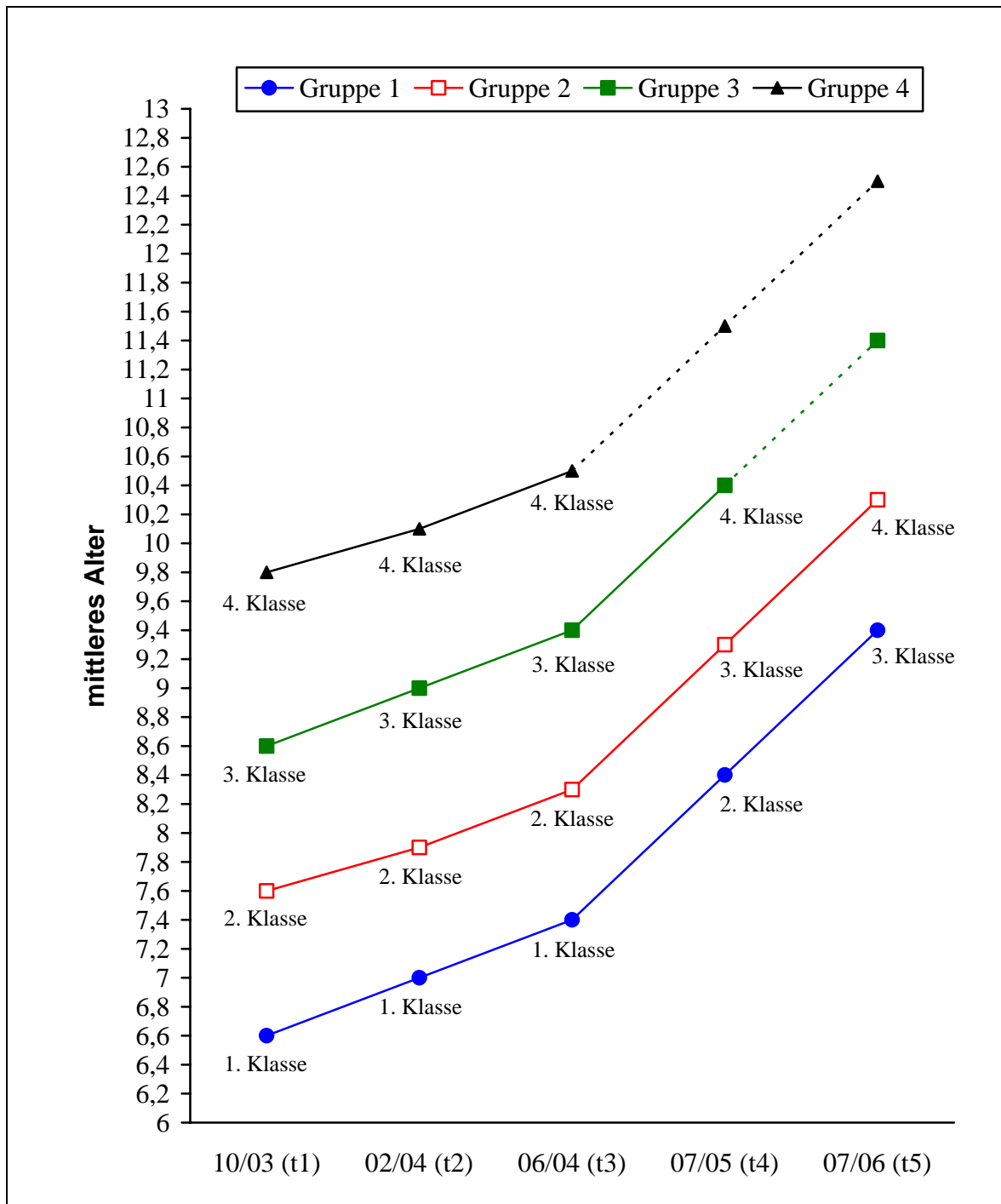


Abbildung 5.1. Mittleres Alter der Schülerinnen und Schüler in den einzelnen Gruppen zu den jeweiligen Testzeitpunkten. Gruppe 1: Erstes Schuljahr zu t1, Gruppe 2: Zweites Schuljahr zu t1, Gruppe 3: Drittes Schuljahr zu t1, Gruppe 4: Viertes Schuljahr zu t1.

Die Geschlechterverteilung unterschied sich in den Schulen nicht signifikant voneinander. In der Grundschule Olewig waren 52,4 Prozent der Kinder Mädchen und 46,4 Jungen; in der Egbert-Schule war das Verhältnis 39,8 Prozent Mädchen respektive 59,0 Prozent Jungen (bei jeweils fehlenden Angaben).

Die Angaben der Kinder zu ihren „Lieblingsfächern“ unterschieden sich nicht wesentlich zwischen den beiden Schulen: Über alle Klassenstufen hinweg wurde Sport als das „beliebteste“ Fach angegeben. Auch hinsichtlich des sozioökonomischen Hintergrunds, der familiären Situation und der Freizeitaktivitäten der Kinder kann aufgrund von Angaben der Eltern oder der Kinder selbst von einer Vergleichbarkeit der beiden Schulen ausgegangen werden.

6 Beschreibung der eingesetzten Testverfahren

6.1 Erfassung des Niveaus der allgemeinen Intelligenz

Das Niveau der allgemeinen Intelligenz wurde mittels des *Grundintelligenztest Skala 1 (CFT 1)* sowie bei den älteren Kindern (ab dem dritten Schuljahr) in der Version *Skala 20 (CFT 20)* erfasst. Bei dem Verfahren handelt es sich um einen sog. „*Culture Fair Test*“ (*CFT*), d.h. dass die Bestimmung der Grundintelligenz ohne Rückgriff auf sprachliches Material resp. sprachgebundene Fertigkeiten möglich sein soll. Erfasst wird die Fähigkeit eines Kindes, anhand von figuralem Material Beziehungen herstellen, Regeln erkennen, Merkmale/Reize identifizieren und verarbeiten zu können. Der Test gibt Aufschluss darüber, bis zu welchem Komplexitätsgrad das Kind in der Lage ist, einzelne dieser sprachfrei präsentierten Aufgaben innerhalb einer für die einzelnen Untertests vorgegebenen Zeitbegrenzung zu erfassen und zu lösen.

Für das Verfahren *CFT 1* liegen Normwerte für den Altersbereich von 5;6 bis 9;5 Jahren vor, es kann auch als Gruppenverfahren angewendet werden. Hierzu gibt es zwei „Pseudo“-Parallelformen, indem die Aufgaben in den beiden Formen identisch und nur die Reihenfolge der Darbietung und die Position der richtigen Lösungen verändert sind. Dadurch soll ein Abschreiben zwischen den Testpersonen erschwert werden. Dieses Verfahren wurde in beiden Formen im ersten und zweiten Schuljahr eingesetzt. Die Summe aller Richtiglösungen stellt das Maß (Rohwerte) der Grundintelligenz dar. Die einzelnen Untertests des *CFT 1* lassen sich wie folgt beschreiben und jeweils anhand eines Beispiels illustrieren:

Untertest 1: Substitution

Vorgegeben sind sechs figurale Vorgaben mit zugehörigen Symbolen. Diese Symbole sollen vom Kind unter die in unterschiedlicher Reihenfolge wiederkehrenden figürlichen Darstellungen gesetzt werden. Erfasst wird die Fähigkeit, schnell die passenden Symbole zu erkennen und den Darstellungen richtig zuzuordnen (=reproduktiver Aspekt der Wahrnehmung; siehe Abbildung 6.1).

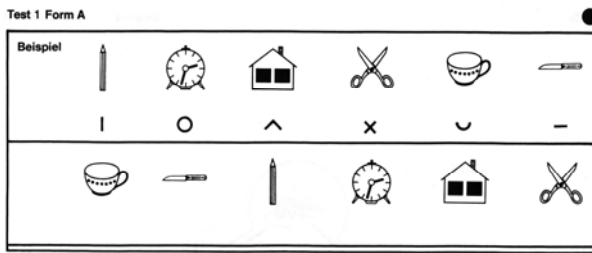


Abbildung 6.1. Beispielaufgaben des CFT 1: Untertest 1 „Substitution“

Untertest 2: Labyrinth

Vier Labyrinth sollen in der vorgegebenen Zeit richtig durchfahren werden. Erfasst werden der optische Wahrnehmungsumfang und die Wahrnehmungsgeschwindigkeit durch Vorgabe einer durch Irrwege erschwerten visuellen Aufgabe (=produktiver Aspekt der Wahrnehmung sowie visuelle Orientierung und Aufmerksamkeit; siehe Abbildung 6.2).

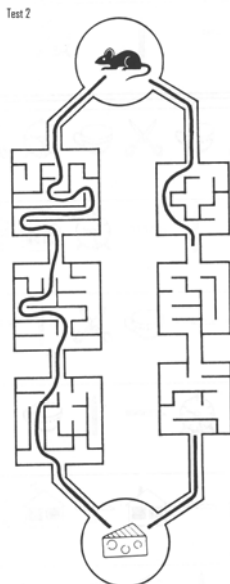


Abbildung 6.2. Beispielaufgaben des CFT 1: Untertest 2 „Labyrinth“

Untertest 3: Klassifikation

Von jeweils fünf figuralen Darstellungen ist immer eine von vier merkmalsähnlichen Figuren abzugrenzen. Erfasst werden die Fähigkeit des Klassifizierens bzw. das beziehungsstiftende Denken bei figuralem Material (siehe Abbildung 6.3).

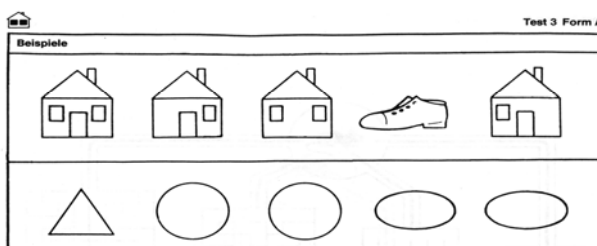


Abbildung 6.3. Beispielaufgaben des CFT 1: Untertest 3 „Klassifikation“

Untertest 4: Ähnlichkeiten

Vorgegebene figurale Darstellungen sollen genau erfasst und aus jeweils fünf weiteren Figuren, von denen vier detail- bzw. merkmalsverändert sind, die mit der Vorgabe identische Figur herausgefunden werden. Erfasst wird, bis zu welchem Komplexitätsgrad figurale Vorgaben wiedererkannt werden, wenn diese im Kontext ähnlicher, jedoch merkmalsveränderter Figuren stehen (siehe Abbildung 6.4).

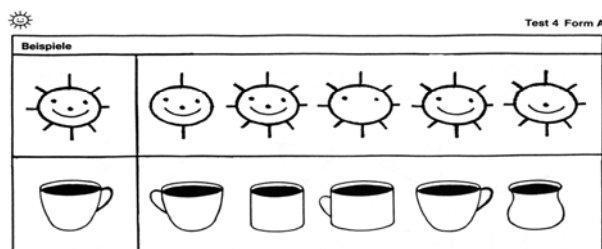


Abbildung 6.4. Beispielaufgaben des CFT 1: Untertest 4 „Ähnlichkeiten“

Untertest 5: Matrizen

Von jeweils fünf figuralen Darstellungen soll diejenige ausgewählt werden, die ein vorgegebenes Muster richtig vervollständigt. Erfasst wird die Fähigkeit, Regeln und Zusammenhänge bei figuralen Problemstellungen zu erkennen (siehe Abbildung 6.5).

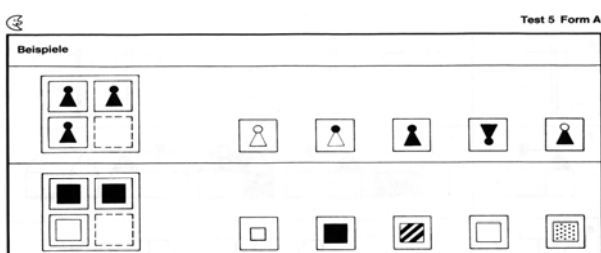


Abbildung 6.5. Beispielaufgaben des CFT 1: Untertest 5 „Matrizen“

Der *Grundintelligenztest Skala 2 (CFT 20)* ist ein Verfahren zur Bestimmung der Grundintelligenz im Altersbereich von 8;7 bis 18;0 Jahren (Normstichprobe) und folgt der gleichen Logik wie der *CFT 1*; es kann auch bei Erwachsenen mit einfacher Schulbildung eingesetzt werden. Wie beim *CFT 1* liegen auch hier zwei Parallelformen vor. Das Verfahren wurde in beiden Formen im dritten und vierten Schuljahr eingesetzt. Der *CFT 20* besteht aus den folgenden vier Subtests:

Untertest 1: Reihenfortsetzen




Die Gesetzmäßigkeiten in einer Reihe figuraler Darstellungen sollen erkannt und die Reihe soll logisch fortgesetzt werden, indem die richtige Lösung aus fünf Antwortalternativen gewählt wird (siehe Abbildung 6.6).

erstes Beispiel




■	■	■		a	b	c	d	e
---	---	---	--	---	---	---	---	---

Bei diesem Beispiel ist **a** die richtige Lösung.
Sie ist auf eurem Antwortbogen bereits angestrichen.

zweites Beispiel

				a	b	c	d	e
--	--	--	--	---	---	---	---	---

drittes Beispiel

				a	b	c	d	e
---	---	---	--	---	---	---	---	---

In jeder Reihe soll also immer diejenige von den 5 Figuren auf der rechten Seite ausgewählt werden, die zu den 3 Figuren auf der linken Seite am besten paßt.

Den Buchstaben, der zu dieser Figur gehört, sollt ihr dann jeweils auf eurem Antwortbogen anstreichen.



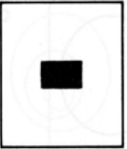


Auf den beiden nächsten Seiten findet ihr 12 Aufgaben, die ihr genauso lösen sollt, wie wir das eben geübt haben.

Abbildung 6.6. Beispielaufgaben des CFT 20: Untertest 1 „Reihenfortsetzen“

Untertest 2: Klassifikation


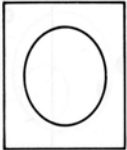
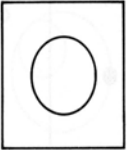

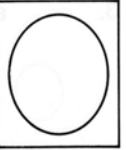
Die Probanden sollen aus einer Vorlage von fünf Figuren eine als von den restlichen (merkmalsähnlichen) Figuren verschieden erkennen und abgrenzen (siehe Abbildung 6.7).

erstes Beispiel

	a	b	c	d	e
					

Beim ersten Beispiel ist **d** die richtige Lösung. Sie ist auf eurem Antwortbogen unter Test 2 bereits angestrichen.

zweites Beispiel

	a	b	c	d	e
					

In jeder Reihe sollt ihr also herausfinden, welche der Figuren in den 5 Kästchen sich von den anderen 4 Figuren in irgendeiner Weise unterscheidet, also nicht zu den anderen Figuren paßt.

Diesmal findet ihr 14 Aufgaben, die ihr in der gleichen Weise lösen sollt.

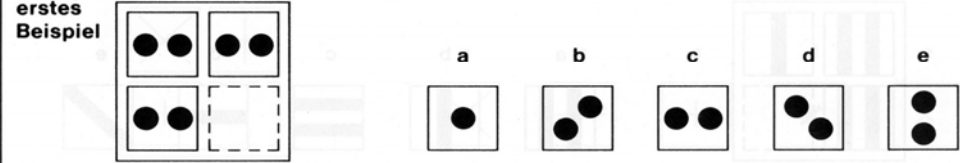
Beachtet bitte, daß die Aufgaben wieder auf 2 Seiten stehen. Wenn ihr mit der ersten Seite fertig seid, arbeitet gleich auf der nächsten Seite weiter.

Abbildung 6.7. Beispielaufgaben des CFT 20: Untertest 2 „Klassifikation“

Untertest 3: Matrizen

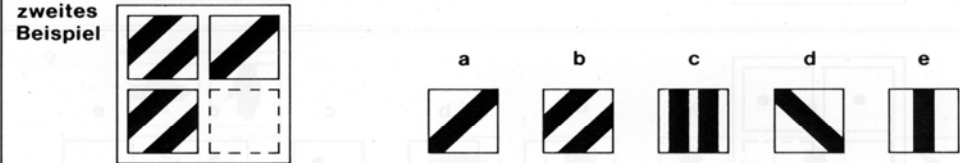
Es soll diejenige von fünf figuralen Darstellungen ausgewählt werden, die ein vorgegebenes Muster richtig vervollständigt. Erfasst wird die Fähigkeit, Regeln und Zusammenhänge bei figuralen Problemstellungen zu erkennen (siehe Abbildung 6.8).

erstes Beispiel

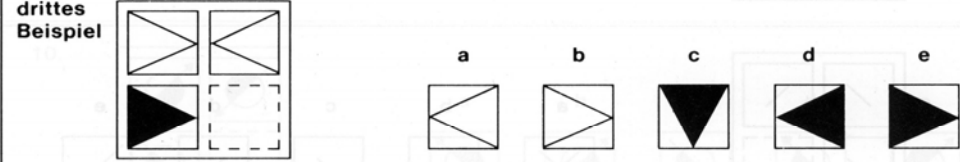


Beim ersten Beispiel ist **c** die richtige Lösung. Diese ist auf dem Antwortbogen unter Test 3 schon angestrichen.

zweites Beispiel



drittes Beispiel



Bei jeder Aufgabe sollt ihr also rechts ein Kästchen mit der Zeichnung auswählen, die in das leere Kästchen links am besten hineinpaßt, um den Kasten richtig zu vervollständigen.

Diesmal sind es wieder 12 Aufgaben, die ihr auf den beiden nächsten Seiten auf diese Weise lösen sollt.

Abbildung 6.8. Beispielaufgaben des CFT 20: Untertest 3 „Matrizen“

Untertest 4: Topologisches Schlussfolgern

Aus fünf präsentierten Darstellungen soll diejenige gewählt werden, in die ein Punkt so gesetzt werden kann, dass dessen Lage der Lage des Punktes in einer vorgegebenen figuralen Darstellung entspricht. Dazu muss die Lage der einzelnen Elemente zueinander in der Vorlage erkannt und auf die Antwortalternativen übertragen werden (siehe Abbildung 6.9).



















erstes Beispiel					
	a 	b 	c 	d 	e 
<p>Beim ersten Beispiel ist c die richtige Lösung, weil man nur hier den Punkt in den Kreis setzen kann, ohne daß er gleichzeitig im Viereck liegt. Auf eurem Antwortbogen ist deshalb der Buchstabe c unter Test 4 bereits angestrichen.</p> <p>Bei den folgenden Beispielen und bei den Aufgaben dürft ihr den Punkt aber nicht einzeichnen, sondern nur den entsprechenden Buchstaben auf eurem Antwortbogen anstreichen.</p>					
zweites Beispiel					
	a 	b 	c 	d 	e 
drittes Beispiel					
	a 	b 	c 	d 	e 
<p>Bei jeder Aufgabe sollt ihr also zunächst genau hinsehen, wo der Punkt liegt und dann unter den 5 Auswahlfiguren diejenige herausuchen, in der der Punkt genauso liegen könnte. Den Lösungsbuchstaben tragt ihr dann auf eurem Antwortbogen ein. Diesmal sind nur 8 Aufgaben zu bearbeiten.</p> <p>Denkt aber daran, daß ihr nichts in das Heft zeichnen dürft, also bitte keinen Punkt in die Aufgaben einzeichnen!</p>					

Abbildung 6.9. Beispielaufgaben des CFT 20: Untertest 4 „Topologisches Schlussfolgern“

6.2 Erfassung der Konzentrationsfähigkeit

Die Konzentrationsfähigkeit wurde mit zwei unterschiedlichen Verfahren erfasst: (1) mit dem *Differentielle Leistungstest (DL-KG und DL-KE)* sowie (2) der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)*.

Der *Differentielle Leistungstest* ist ein Test zur Erfassung des Lernverhaltens bei konzentrierter Tätigkeit im Grundschulalter und liefert Maße für die Konzentrationsfähigkeit von Schülerinnen und Schülern. Er ist normiert für den Altersbereich von sechs bis zehn Jahren. Der Test liegt in zwei Formen vor: Die Form *DL-KE* ist für das erste Schuljahr geeignet, die Form *DL-KG* eignet sich für die Schuljahre (2) bis (4). Die Aufgabe der Schülerinnen und Schüler besteht darin, relevante Reize von irrelevanten Reizen unter Zeitdruck zu unterscheiden. In einer Serie von Abbildungen konkreter und vertrauter Gegenstände (= Reizvorlage, siehe Abbildung 6.10) werden einige als „relevante“ Reize definiert, die über die Vorlage hinweg durchzustreichen sind, während die anderen als „irrelevant“ definierten Reize mit einem Punkt zu versehen sind. In Abhängigkeit vom Altersbereich der Testperson variieren die Vorlage und die Anzahl der zu bearbeitenden Zeichen sowie die Testzeit. Hinzu kommt, dass die Testpersonen alle 90 Sekunden durch den Testleiter aufgefordert werden, das gerade bearbeitete Zeichen zusätzlich mit einem Kreis zu versehen, so dass die einzelnen Bearbeitungsintervalle kenntlich gemacht werden (siehe Abbildungen 6.11, 6.12 und 6.13). Die Auswertung liefert einen „quantitativen“ und einen „qualitativen“ Wert; ersterer bezieht sich auf die Gesamtzahl der angekreuzten Zeichen; der zweite bezieht sich auf die Fehlergesamtzahl relativiert an der Gesamtzahl bearbeiteter Zeichen. Der Wertebereich schwankt zwischen 0 und 98 für den DL-KE sowie zwischen 0 und 240 für den DL-KG.

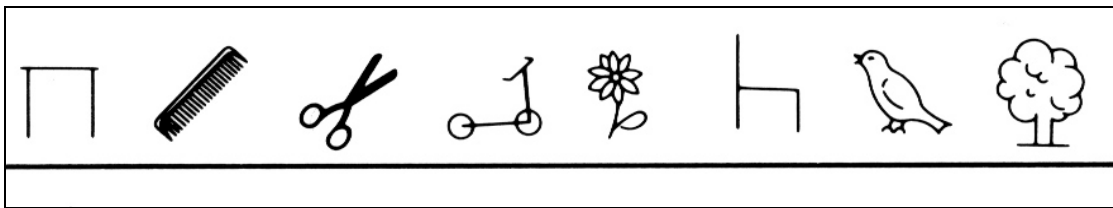


Abbildung 6.10. Auszug aus der Reizvorlage des DL-KE/KG

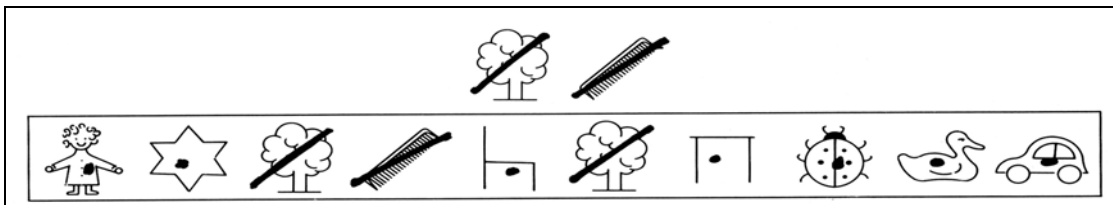


Abbildung 6.11. Abbildung der relevanten Reize und Darstellung eines korrekt bearbeiteten Testabschnitts

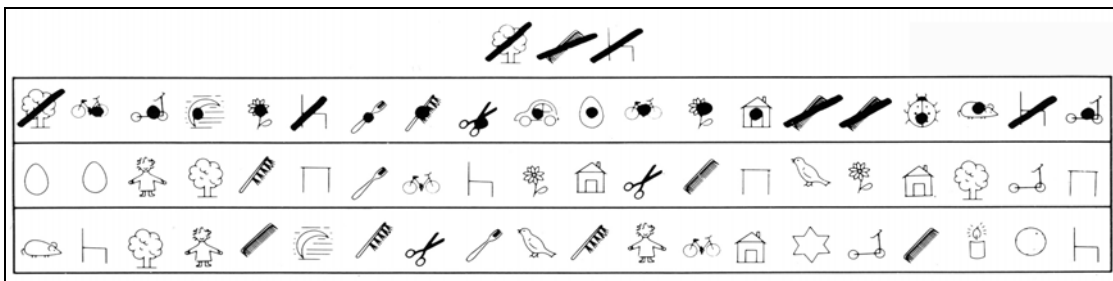


Abbildung 6.12. Abbildung der relevanten Reize und Darstellung eines korrekt bearbeiteten Testabschnitts

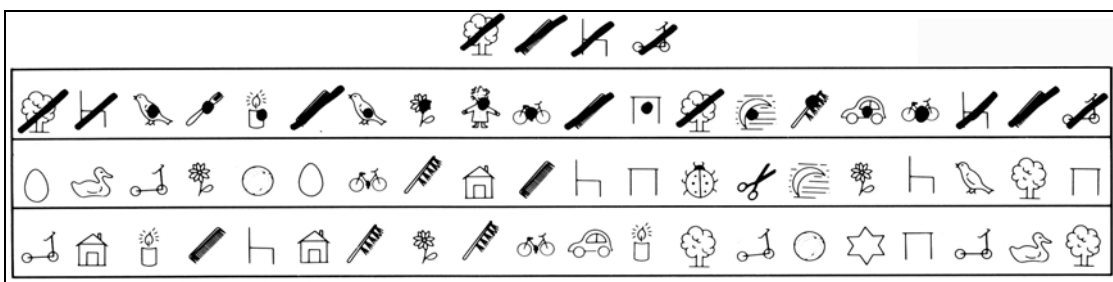


Abbildung 6.13. Abbildung der relevanten Reize und Darstellung eines korrekt bearbeiteten Testabschnitts

Die *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* ermöglicht eine differenzierte Diagnostik von Konzentrationsstörungen bei Schülerinnen und Schülern der Klassen (2) bis (6). Das Verfahren besteht aus einer dreiteiligen Aufgabenserie mit den unten genannten Bestandteilen; es erlaubt auch die Bestimmung von fünf Messwerten, mit denen sowohl die Leistungsmenge wie auch die Genauigkeit der Bearbeitung geprüft werden können:

- Abschreibtext: Hier sind ein bis zwei von Bildern umrahmte Geschichten mit einer Zeitbeschränkung von zehn Minuten abzuschreiben; ermittelt werden die Gesamtzahl der geschriebenen Silben und die Fehlerzahl.
- Rezeptive Aufmerksamkeitsprobe: Hier soll auf die in einer vorgelesenen Geschichte genannten Tierarten geachtet werden, die anschließend aufzuschreiben sind; ermittelt wird die Anzahl der richtig genannten Tierarten.
- Rechentest: Dieser besteht aus gemischten Additions- und Subtraktionsaufgaben, die binnen zehn Minuten zu bearbeiten sind; ermittelt wird die Gesamtzahl der bearbeiteten Aufgaben sowie die Fehlerzahl.

6.3 Erfassung der Rechtschreibfähigkeit

Hierzu wurde die *Hamburger Schreib-Probe für 1. bis 9. Klasse (HSP 1-9)* zur Erfassung der Rechtschreibfähigkeit eingesetzt. Es ist normiert für Schüler und Schülerinnen im Grundschulalter sowie in der Sekundarstufe I und erlaubt sowohl die Einschätzung individueller als auch klassenbezogener Leistungsniveaus; es kann als Einzel- wie als Gruppentest durchgeführt werden. Das Verfahren besteht aus einzelnen Wörtern und Sätzen, die den Kindern vorgelesen werden und die sie anschließend in ihrem individuellen Tempo niederschreiben sollen. Unterstützt werden soll dies durch die zusätzliche Vorgabe von Bildern, die die zu schreibenden Wörter darstellen und die ggf. dabei helfen sollen, dass die Kinder sich an die vorgelesenen Wörter erinnern können. Sollte dennoch ein Wort vergessen werden, so können die Kinder nachfragen. Der bei Diktaten übliche Zwang, einem vorgegebenen Tempo zu folgen, der sowohl für leistungsstärkere wie leistungsschwächere Schülerinnen und Schüler häufig eine Belastung darstellt, entfällt mit der Realisierung dieses Vorgehens. Die Rechtschreibleistung

der Kinder, die anhand des von ihnen geschriebenen Materials bestimmt wird, kann durch verschiedene Parameter abgebildet werden. Im Rahmen der vorliegenden Erhebung wurde die „Anzahl richtig geschriebener Wörter“ als Leistungsmaß gewählt.

6.4 Aspekte der erlebten Integration in die Schule

Neben den genannten Leistungsvariablen sollten schließlich auch noch ausgewählte affektiv-motivationale Variablen bei den Schülerinnen und Schülern erfasst werden. Hier zu wurde der *Fragebogen zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6)* eingesetzt. Dieser dient der strukturierten Erhebung von Informationen dazu, als wie gut integriert in ihre Schulklasse sich Schüler und Schülerinnen selbst erleben. Es ist also ein *self-report*-Verfahren, d.h. es enthält selbstbezogene Aussagen, zu denen die Schülerinnen und Schüler den Grad des Zutreffens auf ihre eigene Person anhand einer vierstufigen Skala (*stimmt genau, stimmt ziemlich, stimmt weniger und stimmt gar nicht*) angeben sollen. In dem Itempool werden verschiedene Facetten der erlebten Integration in die Schule unterschieden:

- das Ausmaß der „sozialen Integration“, das dadurch bestimmt ist, als wie positiv ein Kind seine Beziehungen zu seinen Mitschülern erlebt (z.B. „In meiner Klasse fühle ich mich allein.“; „Mit meinen Mitschülern vertrage ich mich gut.“).
- das Ausmaß der „emotionalen Integration“, welches die Höhe des emotionalen Wohlbefindens eines Kindes in seiner schulischen Umwelt umschreibt; dieses gilt als einer der zuverlässigsten Indikatoren für schulische Integration insgesamt (z.B. „Ohne Schule wäre alles besser.“; „Mir gefällt es in der Schule sehr gut.“).
- das Ausmaß der „leistungsmotivationalen Integration“, welche dadurch definiert ist, wie gut ein Kind sich den in der Schule gestellten Aufgaben gewachsen fühlt und wie motiviert es demzufolge ist, diese Aufgaben in Angriff zu nehmen. Indirekt werden damit auch eine Selbsteinschätzung der eigenen Fähigkeiten resp. schulisches Selbstvertrauen erfasst (z.B. „Ich kann viele Schulaufgaben nicht.“; „In der Schule kann ich alles.“).

Die Gewinnung der oben genannten Daten in beiden Schulen zu den verschiedenen Zeitpunkten erfolgt im Rahmen von Gruppenerhebungen in den Klassen; die Durchführung oblag studentischen Mitarbeiterinnen und Mitarbeitern der Universität Trier.

7 Auswertungshinweise

Bei der Auswertung der Daten wurde mit Hilfe statistischer Verfahren geprüft, inwieweit die Mittelwerte der betrachteten Schulklassen in der "Experimentalschule" sich von denen der Vergleichsschule ("Kontrollschule") unterscheiden. Sollten systematische und signifikante Unterschiede zugunsten der Experimentalschule erkennbar werden, so könnte dies als Hinweis auf die förderlichen Effekte des Schachunterrichtes in der Grundschule Olewig (Experimentalschule) gedeutet werden. Die verwendeten statistischen Analysen erlauben es zu prüfen, mit welcher Wahrscheinlichkeit Unterschiede zwischen den Werten in den beiden Schulen auch durch zufällige Schwankungen in der Leistung der Schülerinnen und Schüler zustande gekommen sein könnten. Nur wenn diese Wahrscheinlichkeit sehr klein ist (geringer als fünf Prozent), spricht man von einem signifikanten oder statistisch bedeutsamen Unterschied. Wenn sich ein solcher Unterschied zeigt, so kann davon ausgegangen werden, dass sich die untersuchten Kinder tatsächlich in ihrer Leistungsfähigkeit unterscheiden. Wenn weiter davon ausgegangen werden kann, dass sich die Schülerinnen und Schüler der beiden Schulen lediglich darin systematisch unterscheiden, dass die einen Schulschachunterricht erhielten, die anderen jedoch nicht und die Kinder im Hinblick auf die übrigen Lebensbereiche (z. B. sozioökonomischer Status der Familie, Bildungsniveau der Eltern) vergleichbar sind, so können die ermittelten Unterschiede, sofern sie zugunsten der Grundschule Olewig ausfallen, als Hinweis auf die positive Wirkung des Schulschachunterrichts gewertet werden.

8 Ergebnisse

Bei der Auswertung der Daten wurden im Wesentlichen zwei Arten von Vergleichen angestellt. Zum einen wurde geprüft, inwieweit zu jedem der einzelnen Messzeitpunkten statistisch bedeutsame Unterschiede zwischen der Experimental- und der Kontrollschule vorliegen. Um auf einen positiven Einfluss des Schachspiels auf die betrachteten Variablen schließen zu können, müssten sich solche Unterschiede zugunsten der Experimentalschule (Olewig) nachweisen lassen. Zum anderen wurde untersucht, wie sich die Leistungen der Schülerinnen und Schüler einer Klasse über die Zeit hinweg verändern, d.h. ob es zu einem signifikanten Anstieg oder ggf. auch Abfall der Leistung in den verwendeten Testverfahren kommt. Die Analyse dieser längsschnittlichen Entwicklung wurde durchgeführt, um die Entwicklungsverläufe in den betrachteten Leistungsbereichen bei Schülerinnen und Schülern mit *und* ohne Schachunterricht dokumentieren zu können. Indem zusätzlich geprüft wurde, ob sich diese Entwicklungsverläufe zwischen den einzelnen Messzeitpunkten in den beiden Schulen voneinander unterscheiden, ist es möglich, den im Grundschulalter grundsätzlich zu erwartenden altersgebundenen Entwicklungsfortschritt (wie er sich in der Kontrollschule zeigen sollte) von einem möglichen, die kognitive Entwicklung fördernden, Einfluss des Schachspiels (wie er sich in der Experimentalschule darstellen sollte) zu trennen.

Die gerade beschriebenen Vergleiche wurden zum einen für die Gesamtstichproben der einzelnen Gruppen angestellt. Darüber hinaus wurden für die *Grundintelligenztestskala*, den *Differentiellen Leistungstest*, die *Hamburger Schreibprobe* und den *Fragebogen zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6)* innerhalb der untersuchten Gruppen zwei Teilstichproben von Schülerinnen und Schülern betrachtet, die bei der ersten Erhebung des jeweiligen Verfahrens hohe resp. niedrige Werte im Vergleich mit ihren Klassenkameraden aufwiesen. Mit Hilfe dieser differenzierteren Analyse sollte geprüft werden, ob der Einfluss des Schachunterrichtes von der ursprünglichen Leistungsstärke der Schülerinnen und Schüler in den betrachteten Bereichen abhängt, d.h. ob womöglich gerade schwächere Schüler durch Schachunterricht besonders gut gefördert werden können.

Die Darstellung der Ergebnisse erfolgt getrennt nach den einzelnen Gruppen/ Schulklassen. Zuerst werden die Befunde für Gruppe 1 berichtet, d.h. für diejenigen Schülerinnen und Schüler, die zu Beginn der Evaluationsstudie das erste Schuljahr besuchten. Anschließend werden die Befunde für die Gruppen 2, 3 und 4 in der gerade genannten Reihenfolge beschrieben. Die Ergebnisdarstellung umfasst Textmaterial sowie eine Serie von Abbildungen; ergänzend können im Anhang des Berichts Tabellen mit weiteren Informationen eingesehen werden. Innerhalb jeder Gruppe werden in jeweils gleicher Reihenfolge die Ergebnisse zu den einzelnen Testverfahren dargestellt. In den Abbildungen stehen fett gedruckte Linien für eine signifikante Veränderung der Leistung in der entsprechenden Schule von dem betreffenden Messzeitpunkt zum nächsten. Sterne bezeichnen statistisch bedeutsame Leistungsunterschiede zwischen Experimentalschule und Kontrollschule zu dem betreffenden Messzeitpunkt. Dabei bezeichnet ein Stern (*), dass die Wahrscheinlichkeit, dass der gefundene Unterschied durch Zufall zustande gekommen sein könnte, geringer ist als fünf Prozent; bei zwei Sternen (**) ist diese Wahrscheinlichkeit sogar geringer als ein Prozent.

8.1 Befunde für Gruppe 1

Die Schülerinnen und Schüler der Gruppe 1 besuchten zu Beginn der Evaluationsstudie das erste Schuljahr. Zum vorläufig letzten Messzeitpunkt im Juli 2006 (t_5) waren diese Kinder am Ende ihres dritten Schuljahrs. Folglich hatten diese Kinder in der Experimentalschule (Olewig) bis zu diesem Zeitpunkt drei Jahre lang Schachunterricht in der Schule erhalten.

8.1.1 Niveau der allgemeinen Intelligenz

Die Testleistungen der Kinder beider Schulen in der *Grundintelligenztestskala 1 (CFT 1)* nehmen vom 1. bis zum 3. Messzeitpunkt signifikant zu. Während sich dieser Zuwachs in der Kontrollschule auch noch bis zum 4. Messzeitpunkt substantiell fortsetzt, ist der Leistungsanstieg für die Experimentalschule nicht mehr signifikant, weshalb der Vergleich der beiden Gruppen signifikant bessere Leistungen der Kontrollschule zu t_4 offenbart. Zu den anderen Messzeitpunkten ist der Unterschied in der Leistung der beiden nicht statistisch bedeutsam. Beide Schulen unterscheiden sich zum letzten Zeitpunkt nicht mehr voneinander, und zwar sowohl bezogen auf den Gesamttest wie auch auf einzelne Untertests. Der zum fünften Messzeitpunkt insgesamt deutlich niedrigere Wert für beide Schulen ist durch die Verwendung einer anderen Testversion (Wechsel der Altersnormgruppe) bedingt (vgl. Abb. 8.1).

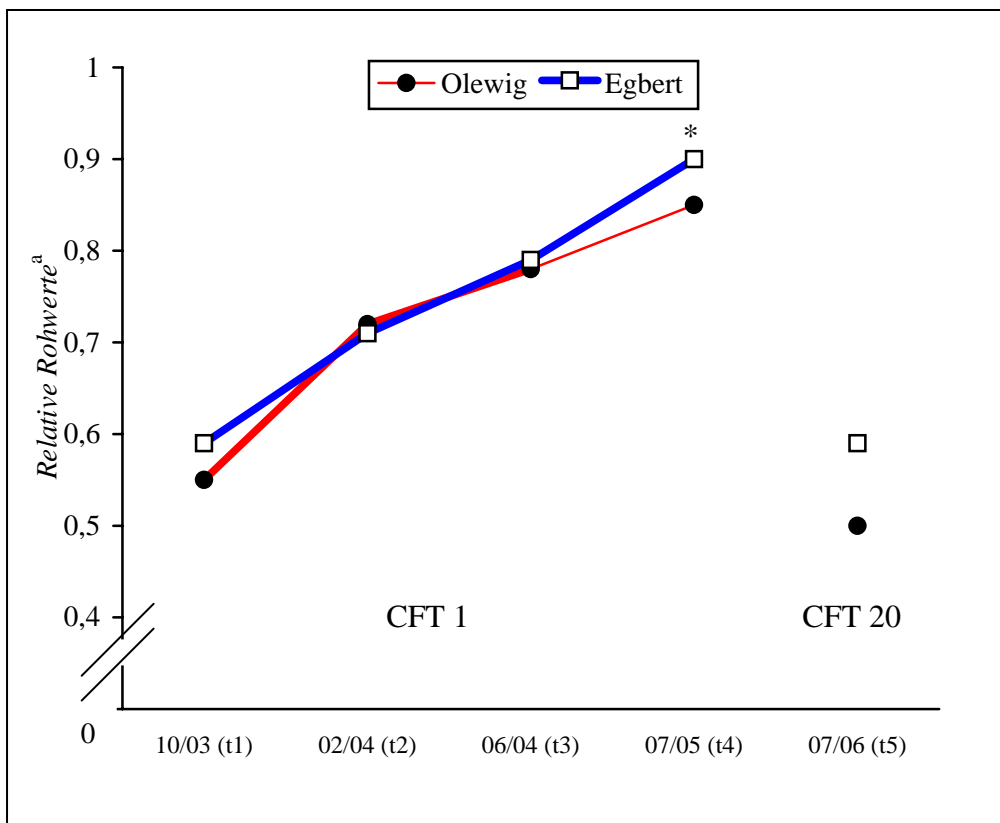


Abbildung 8.1 Relative Rohwerte im CFT1 ($t_1 - t_4$) und CFT20 (t_5) über die fünf Messzeitpunkte für Gruppe 1.

^aDa zum Testzeitpunkt t_5 eine andere Testversion mit einem anderen Wertebereich verwendet wurde, erfolgt die Darstellung aus Gründen der Vergleichbarkeit mit relativen Rohwerten.

In beiden Schulen waren zudem die Schülerinnen und Schüler auf der Grundlage ihrer Leistungen zum ersten Messzeitpunkt in „leistungsschwache“ und „leistungsstarke“ Kinder (jeweils unterhalb oder oberhalb des Medians, also des Wertes, über dem genauso viele Schülerinnen und Schüler einer Klasse liegen wie darunter) unterteilt worden. Für die leistungsschwachen Kinder der Kontrollschule lässt sich ein kontinuierlicher Anstieg vom 1. bis 4. Messzeitpunkt erkennen; in der Experimentalschule findet sich ein signifikanter Anstieg nur von t_1 nach t_2 . Im Vergleich der beiden Schulen zueinander zeigt sich lediglich zu t_4 ein statistisch bedeutsamer Leistungsunterschied zugunsten der Kontrollschule. Für die leistungsstarken Kinder findet sich zu keinem der Messzeitpunkte ein signifikanter Unterschied zwischen den Schulen. Eine bedeutsame Steigerung der Leistung kann nur von t_1 nach t_2 für die Experimentalschule nachgewiesen werden (vgl. Abb. 8.2).

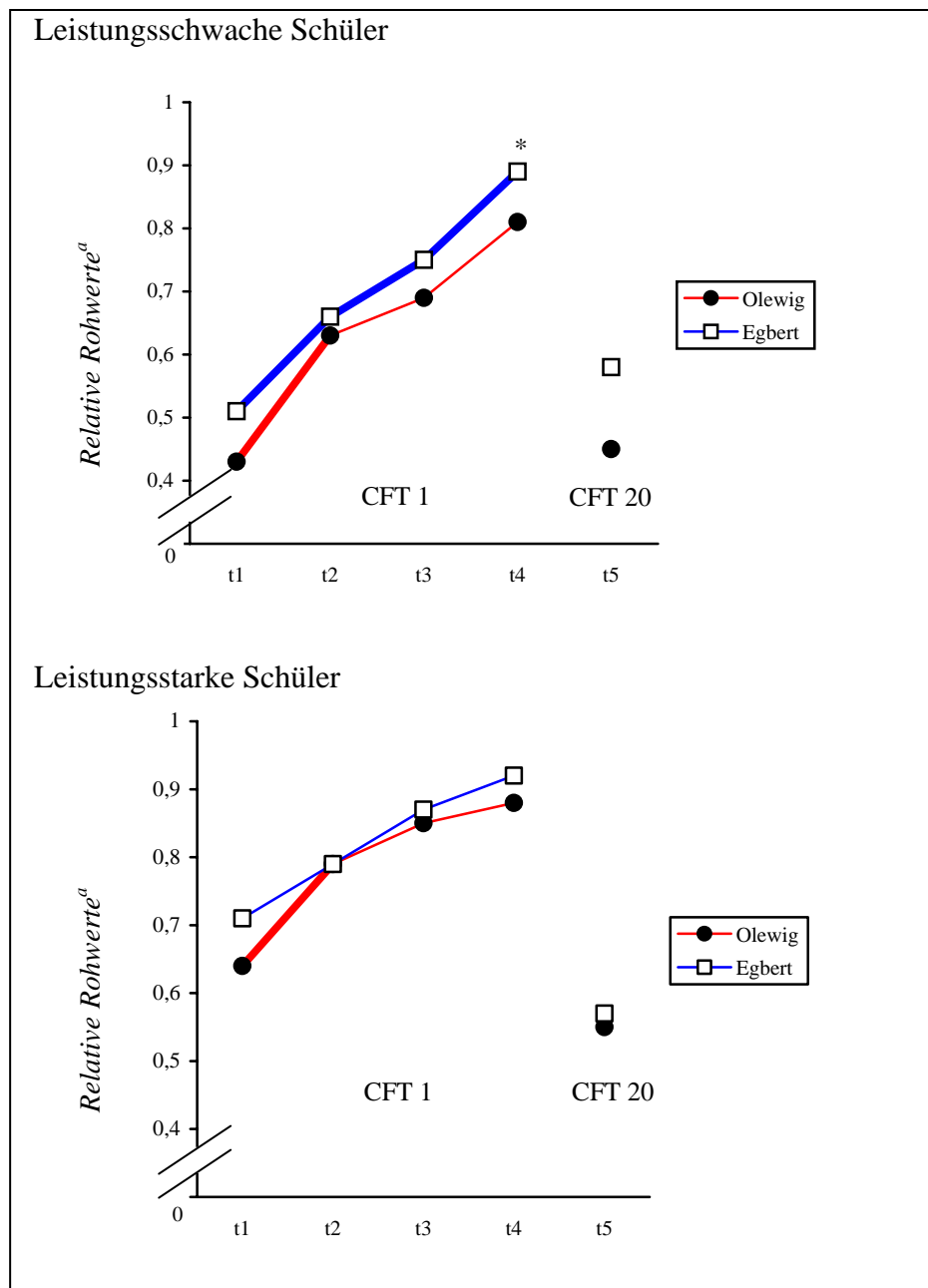


Abbildung 8.2 Gesamtscore der allgemeinen Intelligenz im CFT1 ($t_1 - t_4$) und CFT20 (t_5) über die fünf Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 1.

^aDa zum Testzeitpunkt t_5 eine andere Testversion mit einem anderen Wertebereich verwendet wurde, erfolgt die Darstellung aus Gründen der Vergleichbarkeit mit relativen Rohwerten.

Mit Blick auf die *Untertests* „Substitution“ und „Labyrinth“ (als Maß für produktive und reproduktive Aspekte der Wahrnehmung, visuellen Orientierung und Aufmerksamkeit) steigen die Leistungen der Kinder insgesamt sowie die der leistungsschwachen Kinder in beiden Schulen von t_1 nach t_2 sowie von t_2 nach t_3 jeweils bedeutsam an; für die Kontrollschule zeigt sich noch ein Anstieg von t_3 nach t_4 , nicht aber von t_2 nach t_3 . Unterschiede im Leistungsniveau zwischen den beiden Schulen finden sich nicht (vgl. Abb. 8.3). In der Teilstichprobe der leistungsstarken Kinder finden sich weder ein signifikanter Anstieg noch bedeutsame Unterschiede zwischen den Schulen (vgl. Abb. 8.4).

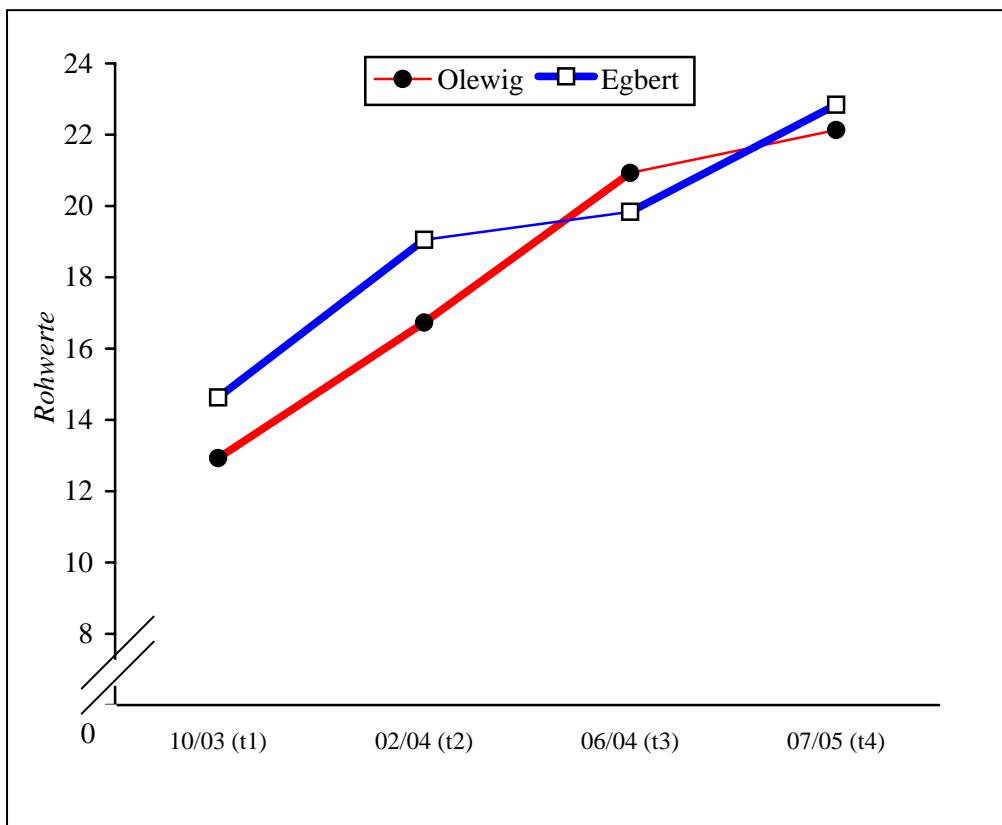


Abbildung 8.3 Summe der Rohwerte in den Untertests „Substitution“ und „Labyrinth“ des CFT 1 über die vier Messzeitpunkte für Gruppe 1 (theoretischer Range: 0 bis 24). Mit dem zu t_5 verwendeten CFT 20 werden keine entsprechenden Werte erfasst.

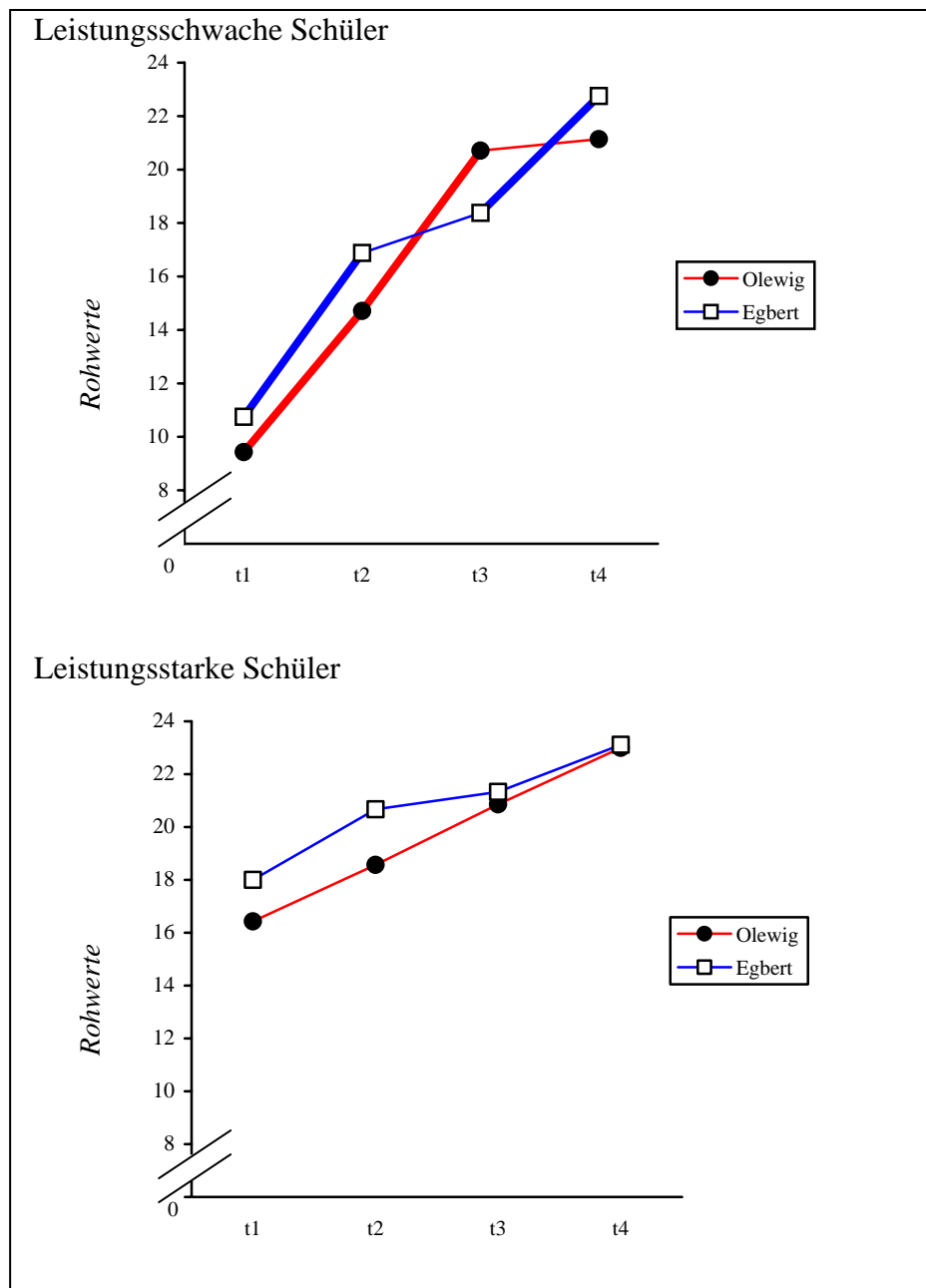


Abbildung 8.4 Summe der Rohwerte in den Untertests „Substitution“ und „Labyrinth“ des CFT 1 über die vier Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 1 (theoretischer Range: 0 bis 24). Mit dem zu t_5 verwendeten CFT 20 werden keine entsprechenden Werte erfasst.

In der Kontrollschule (Egbert) steigt die Leistung in den *Untertests* „Klassifikation“, „Ähnlichkeiten“ und „Matrizen“ (als Maß für beziehungsstiftendes Denken und das Erkennen von Regelmäßigkeiten) kontinuierlich an; in der Experimentalschule (Olewig) findet sich dagegen nur von t_1 nach t_2 ein substantieller Leistungsanstieg, so dass zum vierten Messzeitpunkt die Leistungen der Kontrollschule signifikant über der Leistung der Experimentalschule liegen (vgl. Abb. 8.5).

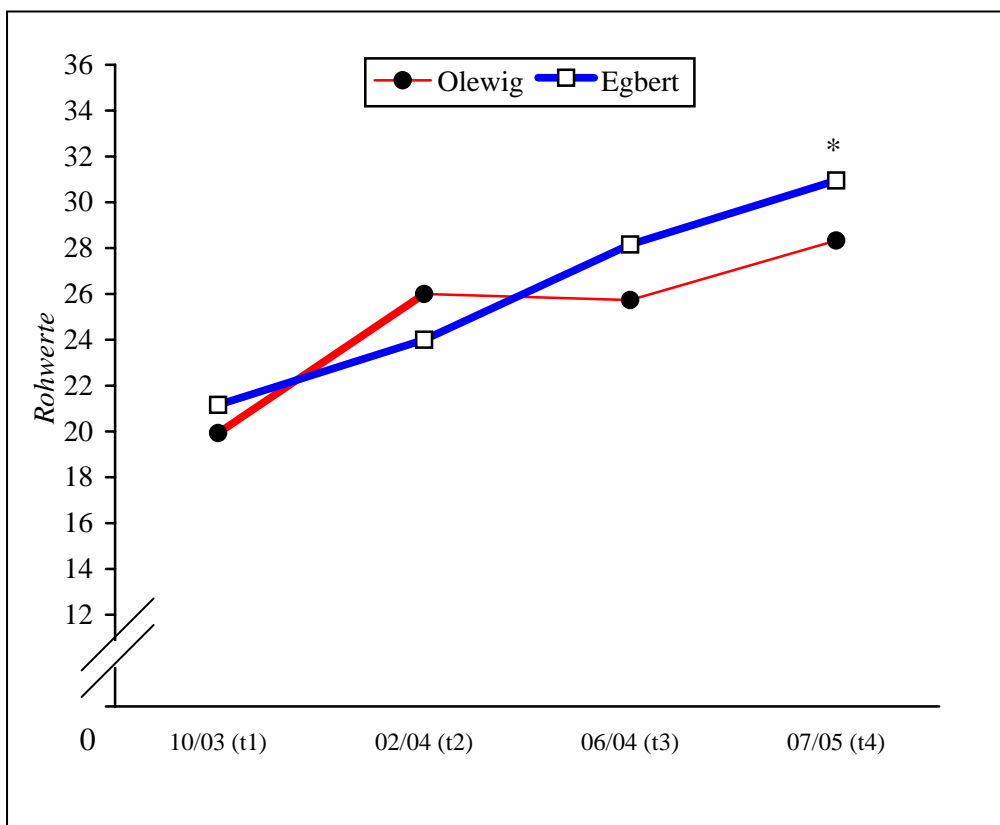


Abbildung 8.5 Summe der Rohwerte in den Untertests „Klassifikation“, „Ähnlichkeiten“ und „Matrizen“ des CFT1 über die vier Messzeitpunkte für Gruppe 1 (theoretischer Range: 0 bis 36). Mit dem zu t5 verwendeten CFT 20 werden keine entsprechenden Werte erfasst.

In der Gruppe der leistungsschwachen Schülerinnen und Schüler steigen die Leistungen der Kinder der Kontrollschule von t_2 nach t_3 und von t_3 nach t_4 , in der Experimentalschule von t_1 nach t_2 sowie von t_3 nach t_4 bedeutsam an. Zum vierten Messzeitpunkt findet sich in der Gruppe der leistungsstarken Schülerinnen und Schüler sich ein signifikanter Unterschied in diesen drei Untertests zugunsten der Kontrollschule. Eine bedeutsame Leistungssteigerung kann nur für die Experimentalschule von t_1 nach t_2 nachgewiesen werden (vgl. Abb. 8.6). Bezogen

auf den letzten Messzeitpunkt, zu dem der *CFT 20* eingesetzt worden war, zeigten sich weder im Gesamtwert noch in den einzelnen Untertests signifikante Unterschiede.

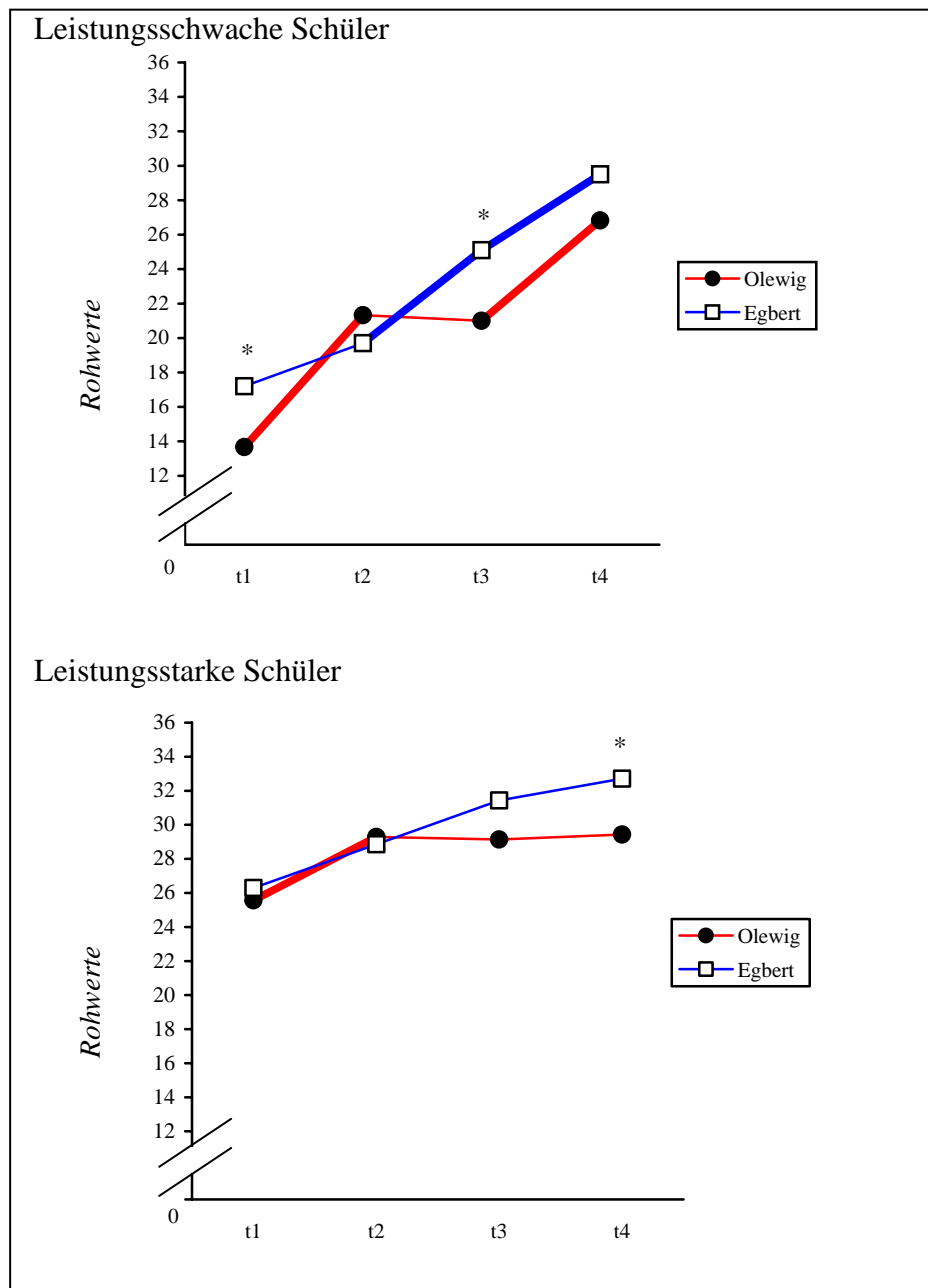


Abbildung 8.6 Summe der Rohwerte in den Untertests „Klassifikation“, „Ähnlichkeiten“ und „Matrizen“ des *CFT1* über die vier Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 1 (theoretischer Range: 0 bis 36). Mit dem zu *t5* verwendeten *CFT 20* werden keine entsprechenden Werte erfasst.

8.1.2 Konzentrationsfähigkeit

Unterschiede in der Höhe der Konzentrationsfähigkeit zwischen den beiden Schulen – erfasst mittels des *Differentiellen Leistungstests (DL-KE / DL-KG)* – bestehen zu keinem der fünf Messzeitpunkte, das gilt auch bei getrennter Betrachtung der leistungsstarken vs. leistungsschwachen Kinder beider Schulen. Die insgesamt niedrigen Werte zu t_4 für beide Schulen sind auf den Einsatz einer anderen Version mit veränderter Metrik zurückzuführen. Gleichwohl ist ein signifikanter Anstieg der Konzentrationsfähigkeit von t_4 nach t_5 bei den Schülerinnen und Schüler mit Schachunterricht (Olewig) erkennbar, der sich für die Schülerinnen und Schüler ohne Schachunterricht (Egbert) nicht zeigt (vgl. Abb. 8.7).

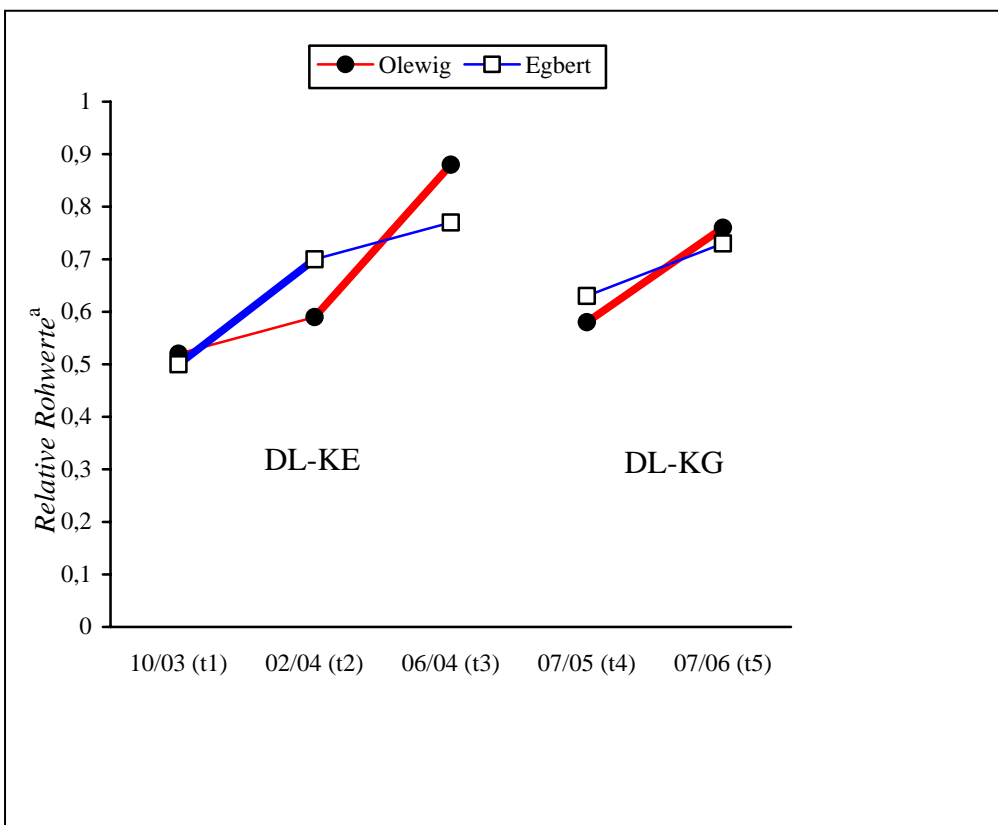


Abbildung 8.7 Relative Rohwerte im DL-KE und DL-KG über die fünf Messzeitpunkte (t_1 - t_5) für Gruppe 1.

^aDa zu den Testzeitpunkten t_4 und t_5 eine andere Testversion mit einem anderen Wertebereich verwendet wurde, erfolgt die Darstellung aus Gründen der Vergleichbarkeit mit relativen Rohwerten.

Bei gesonderter Betrachtung der leistungsschwachen Kinder zeigt sich in der Experimentalschule (Olewig) ein bedeutsamer Leistungsanstieg von t_2 nach t_3 , für die entsprechenden Kinder der Kontrollschule (Egbert) zeigt sich ein solcher allerdings

auch (von t_1 nach t_2). Auch die leistungsstarken Kinder der Grundschule Olewig zeigen einen bedeutsamen Leistungszuwachs von t_2 nach t_3 , während die Kinder der Kontrollschule (Egbert) ihre Konzentrationsleistungen von t_4 nach t_5 bedeutsam steigern können (vgl. Abb. 8.8).

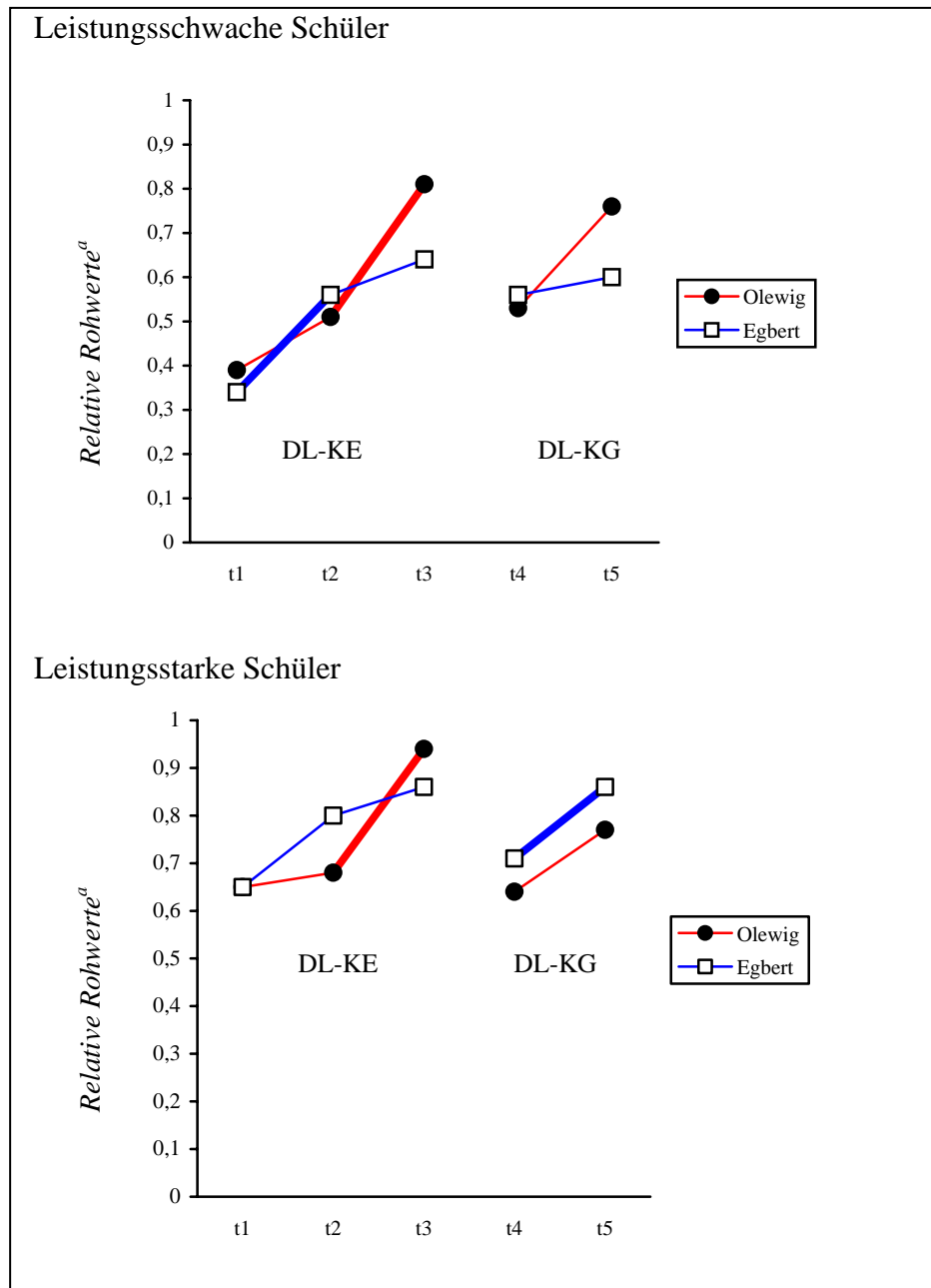


Abbildung 8.8 Relative Rohwerte im DL-KE und DL-KG über die fünf Messzeitpunkte (t_1 - t_5) für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 1.

^aDa zu den Testzeitpunkten t_4 und t_5 eine andere Testversion mit einem anderen Wertebereich verwendet wurde, erfolgt die Darstellung aus Gründen der Vergleichbarkeit mit relativen Rohwerten.

Betrachtet man als Maß der Konzentrationsfähigkeit die Gesamtzahl der abgeschriebenen Silben, wie es die *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* liefert, so zeigt sich in beiden Schulen gleichermaßen ein signifikanter Leistungszuwachs zwischen t_3 und t_4 sowie zwischen t_4 und t_5 , ohne dass sich die Kinder beider Schulen in der Höhe ihrer Konzentrationsfähigkeit zu irgendeinem Messzeitpunkt unterscheiden würden (vgl. Abb. 8.9).

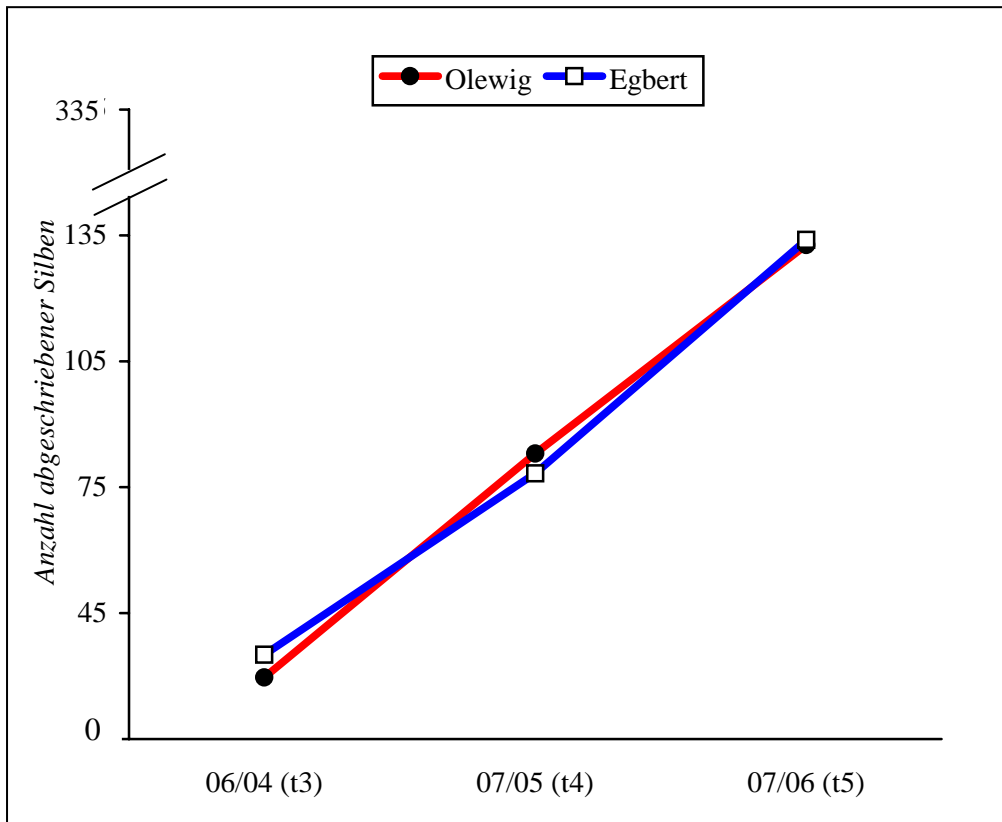


Abbildung 8.9 Anzahl abgeschriebener Silben im TPK über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 1 (theoretischer Range: 0 bis 335).

Für die Fehlerzahl beim Abschreiben der Silben als Maß der Konzentrationsfähigkeit zeigen sich keine bedeutsamen Leistungsveränderungen für beide Schulen über die Zeit und auch keine bedeutsamen Unterschiede zwischen den Schulen zu dem jeweiligen Messzeitpunkt. In der rezeptiven Aufmerksamkeit erzielten die Kinder beider Schulen einen signifikanten, fast gleich verlaufenden Leistungszuwachs von t_3 nach t_4 ; danach zeigt sich für keine der Schulen eine bedeutsame Veränderung, und zu keinem der Messzeitpunkte unterscheiden sich hier die Kinder der Kontrollschule (Egbert) von jenen der Experimentalschule (Olewig, vgl. Abb. 8.10).

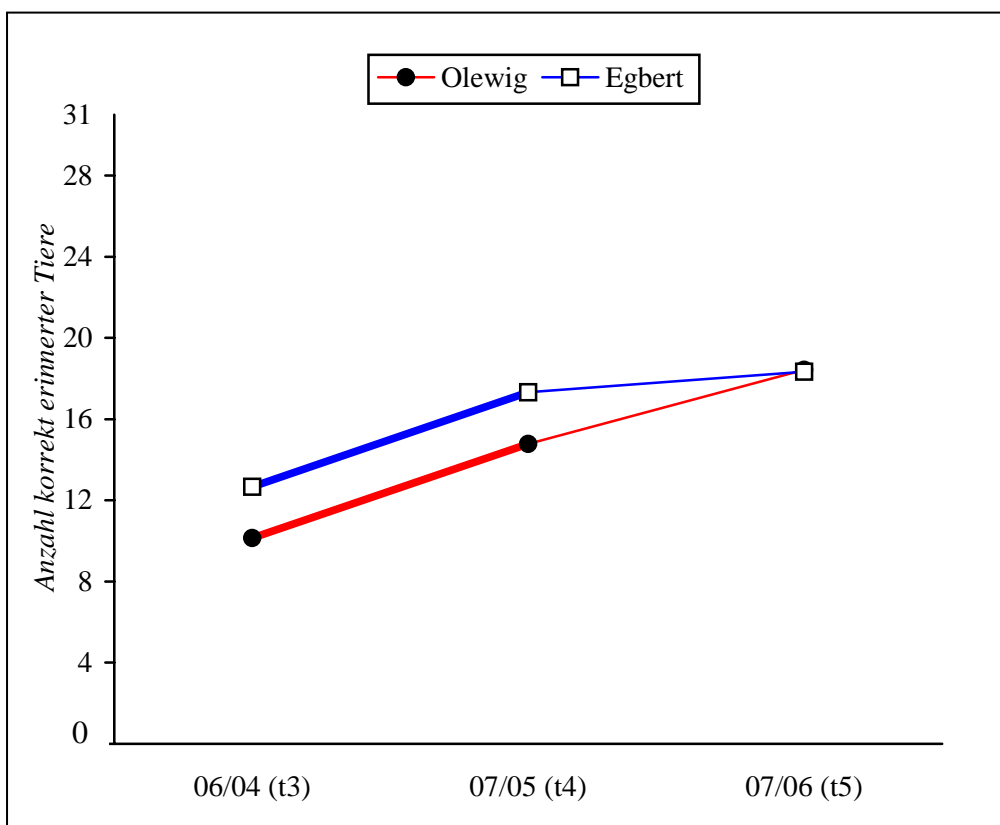


Abbildung 8.10 Anzahl korrekt erinnerter Tiere im TPK über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 1 (theoretischer Range: 0 bis 31).

8.1.3 Prüfung von Rechen- und Rechtschreibleistungen

Mit Blick auf die Gesamtzahl der bearbeiteten Rechenaufgaben, wie sie in der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* vorgegeben werden, findet sich für die Kontrollschule (Egbert) ein signifikanter Leistungsanstieg von t_3 nach t_4 . Von t_4 nach t_5 zeigt sich dagegen ein bedeutsamer Anstieg der Leistung für die Schülerinnen und Schüler mit Schachunterricht (Olewig). Der Vergleich der Schulen zueinander ergibt zu keinem der Messzeitpunkte signifikante Unterschiede (vgl. Abb. 8.11).

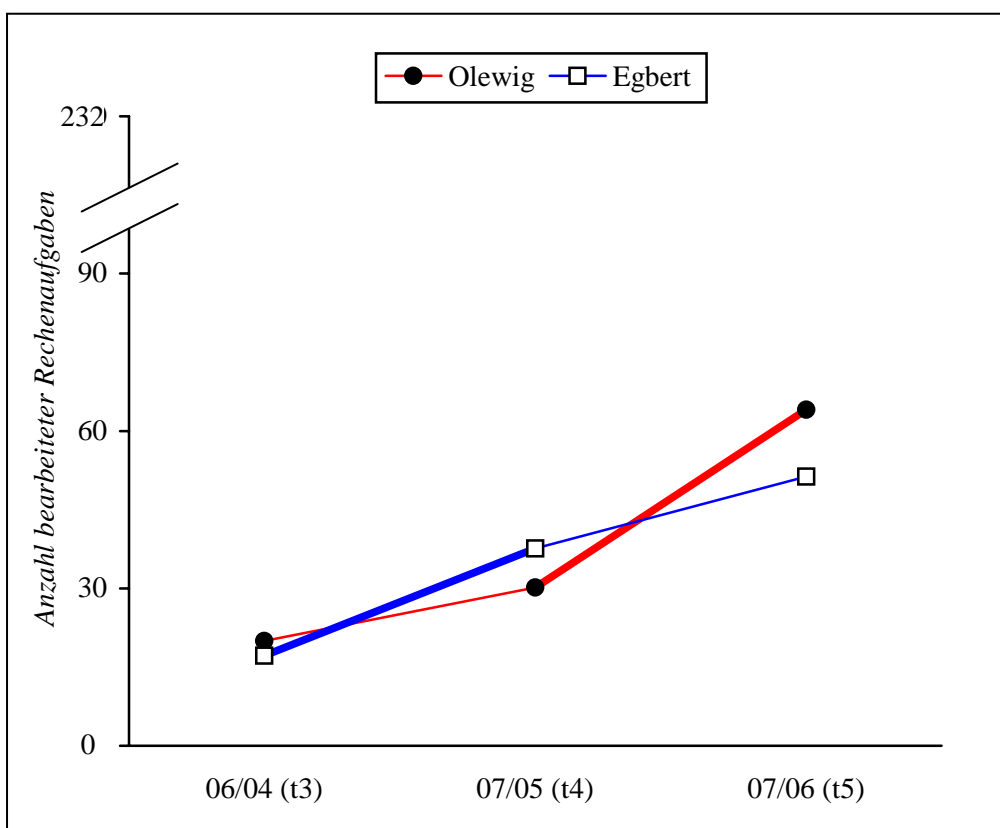


Abbildung 8.11 Anzahl bearbeiteter Rechenaufgaben im TPK über die drei Messzeitpunkte (t_3 – t_5) für Gruppe 1 (theoretischer Range: 0 bis 232).

Für keine der beiden Schulen lässt sich eine bedeutsame Veränderung der Leistung - gemessen über die Fehlerzahl - über die Zeit bzw. die Messzeitpunkte hinweg feststellen. Auch für dieses Maß der Rechenleistung ergibt der Vergleich der beiden Schulen zueinander zu keinem der Messzeitpunkte bedeutsame Unterschiede.

Für die mittels der *Hamburger Schreibprobe (HSP)* erfasste Rechtschreibfähigkeit findet sich ein signifikanter Leistungszuwachs von t_4 nach t_5 nur für die Schülerinnen

und Schüler ohne Schachunterricht (Egbert). Der Vergleich der Schulen zueinander ergibt für keinen der Messzeitpunkte bedeutsame Unterschiede (vgl. Abb. 8.12).

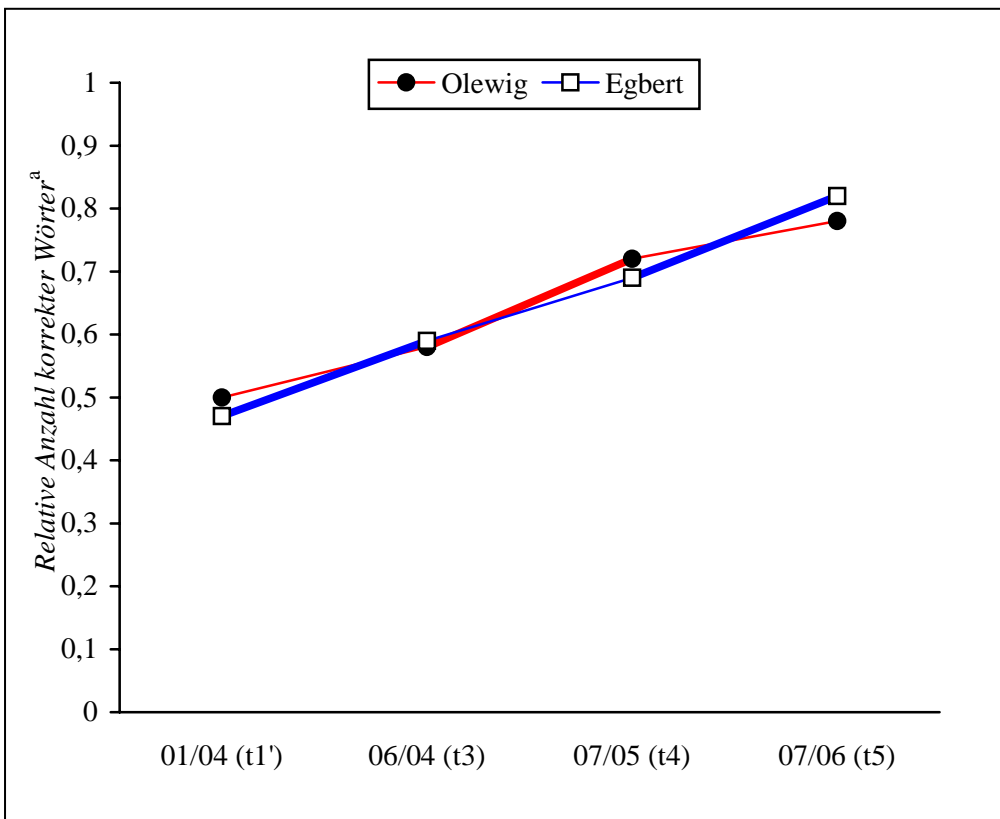


Abbildung 8.12 Relative Anzahl korrekt geschriebener Wörter im HSP über die vier Messzeitpunkte (t_1' – t_5) für Gruppe 1. Die Erhebung zu t_1' wurde verzögert durchgeführt, da zu Beginn des ersten Schuljahres noch keine für eine Testung hinreichenden Rechtschreibfähigkeiten vorausgesetzt werden konnten.

^aDa zu den einzelnen Messzeitpunkten Testhefte mit unterschiedlicher Gesamtzahl von Wörtern verwendet wurden, sind aus Gründen der Vergleichbarkeit die an der Gesamtzahl der Wörter relativierten Häufigkeiten dargestellt.

Die leistungsschwachen Schülerinnen und Schüler der Experimentalschule (Olewig) zeigen einen bedeutsamen Leistungsanstieg von t_1' nach t_3 und von t_3 nach t_4 . In der Kontrollschule (Egbert) zeigt sich ein solcher von t_1' nach t_3 und von t_4 nach t_5 . Bei den leistungsstarken Schülerinnen und Schülern kann lediglich für die Kontrollschule (Egbert) ein signifikanter Leistungszuwachs (von t_4 nach t_5) nachgewiesen werden. Signifikante Unterschiede zwischen den Schulen bestehen nicht (vgl. Abb. 8.13).

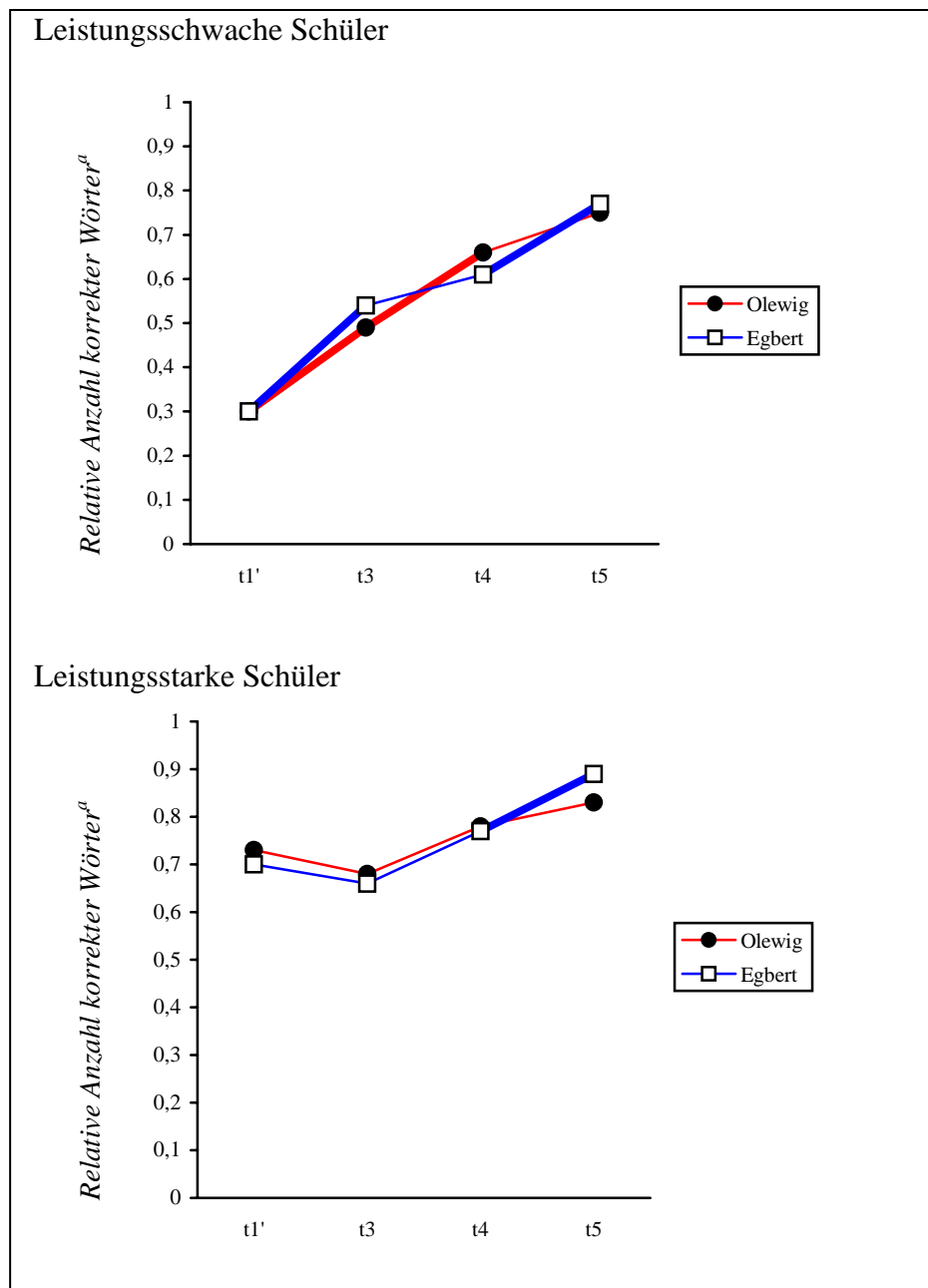


Abbildung 8.13 Relative Anzahl korrekt geschriebener Wörter im HSP über die vier Messzeitpunkte (t_1' – t_5) für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 1. Die Erhebung zu t_1' wurde verzögert durchgeführt, da zu Beginn des ersten Schuljahres noch keine für eine Testung hinreichenden Rechtschreibfähigkeiten vorausgesetzt werden konnten.

^aDa zu den einzelnen Messzeitpunkten Testhefte mit unterschiedlicher Gesamtzahl von Wörtern verwendet wurden, sind aus Gründen der Vergleichbarkeit die an der Gesamtzahl der Wörter relativierten Häufigkeiten dargestellt.

8.1.4 Aspekte der schulischen Integration

Mit Blick auf den *Fragebogen zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6)* zeigen sich für die Dimension „Soziale Integration“ keinerlei Veränderungen über die Zeit bzw. die einzelnen Messzeitpunkte hinweg, und zwar weder in der Experimental- noch in der Kontrollschule. Auch im Vergleich der beiden Schulen zueinander zeigt sich zu keinem Messzeitpunkt ein bedeutsamer Unterschied. Für die „Emotionale Integration“ zeigt sich indes eine bedeutsame Abnahme von t_3 nach t_5 , und zwar in beiden Schulen. Der Vergleich der Schulen zueinander erbringt zu keinem Messzeitpunkt einen signifikanten Unterschied (vgl. Abb. 8.14).

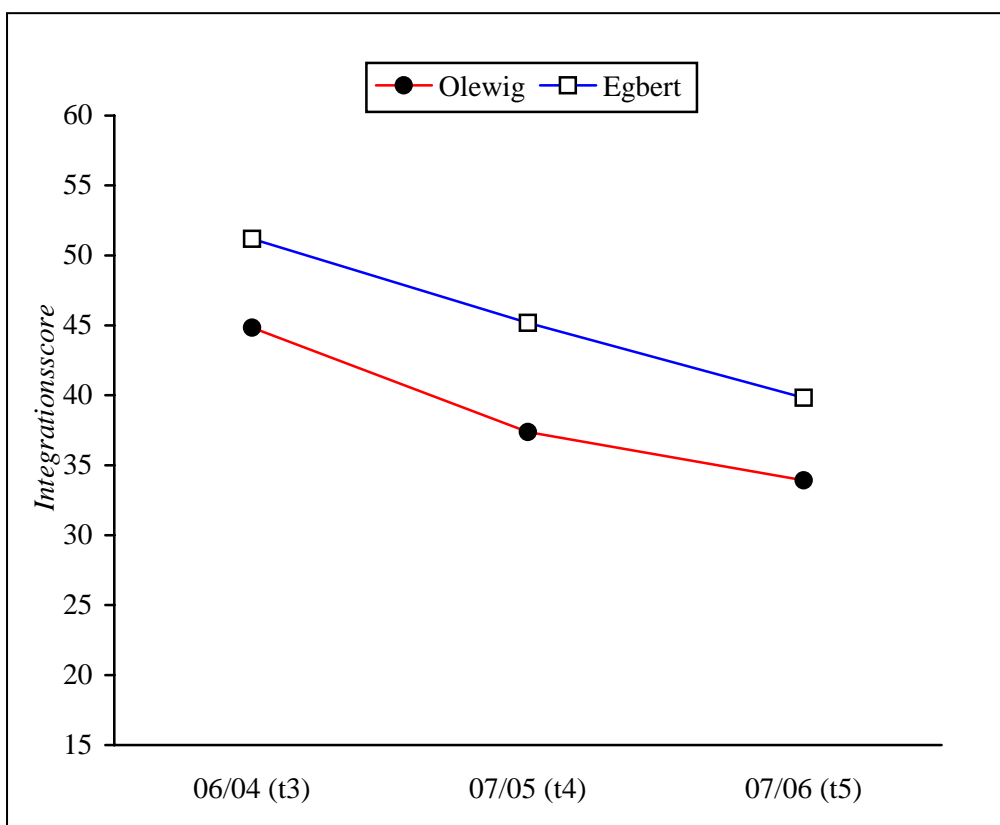


Abbildung 8.14 Scores für emotionale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 1 (theoretischer Range: 15 bis 60).

In der nach der Höhe der emotionalen Integration zu t_3 differenzierten Betrachtung sind weder für die hoch integrierten noch für die schlecht integrierten Schülerinnen und Schüler beider Schulen bedeutsame Veränderungen in der emotionalen Integration zwischen den einzelnen Messzeitpunkten zu beobachten. Die im Vergleich mit ihren Klassenkameraden schlecht integrierten Schülerinnen und Schüler der Kontrollschule (Egbert) weisen jedoch zu allen drei Messzeitpunkten signifikant höhere Werte auf der Dimension „Emotionale Integration“ auf als die schlecht integrierten Schülerinnen und Schüler der Experimentalschule (Olewig). Auf Seiten der Schülerinnen und Schüler mit relativ hohen Werten auf dieser Dimension der schulischen Integration findet sich ein solcher Unterschied nur zu t_3 (vgl. Abb. 8.15).

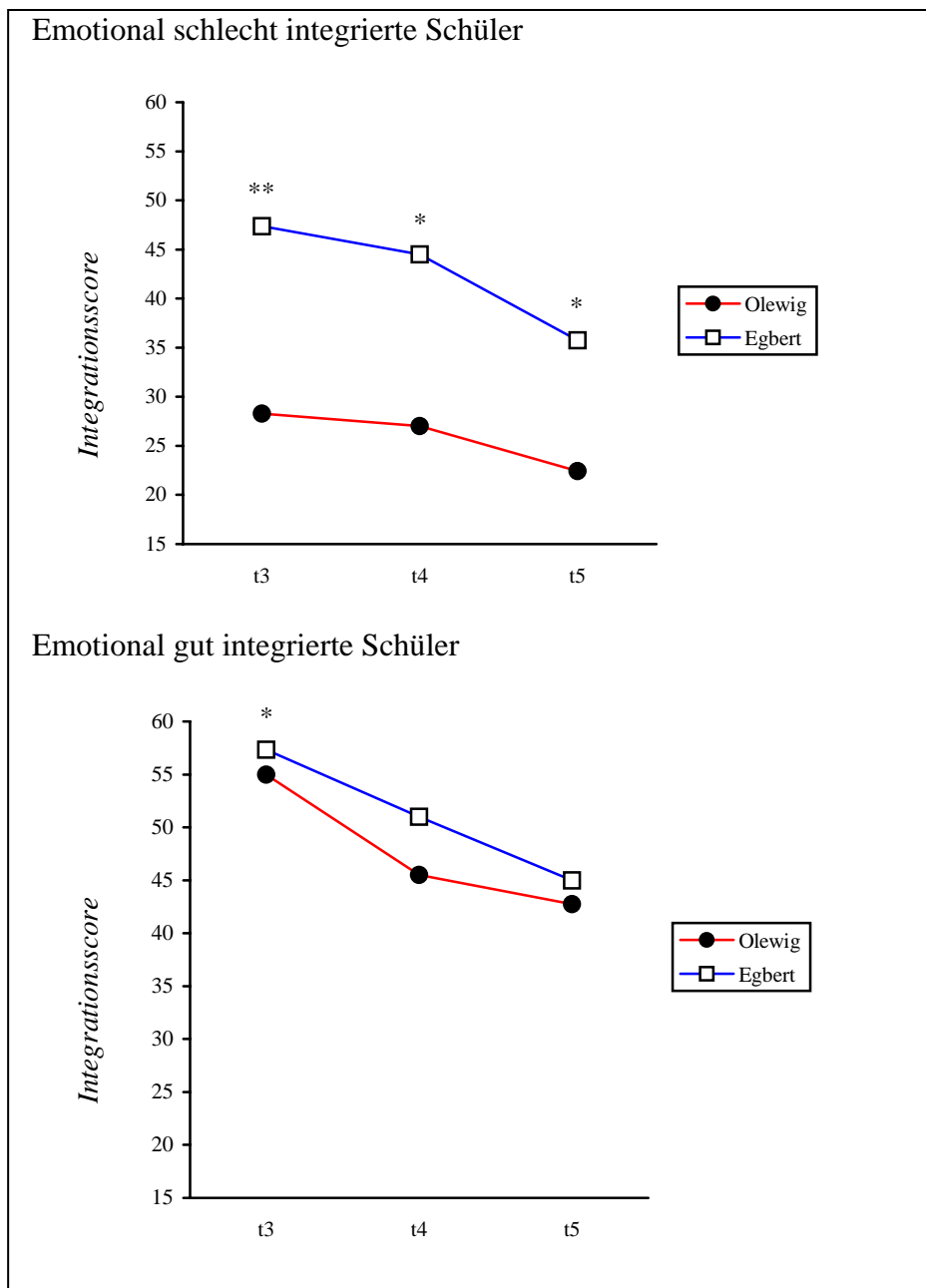


Abbildung 8.15 Scores für emotionale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für schlecht integrierte resp. gut integrierte Schüler aus Gruppe 1 (theoretischer Range: 15 bis 60).

Auf der Dimension „*Leistungsmotivationale Integration*“ ist für keine der beiden Gruppen eine bedeutsame Veränderung über die Zeit bzw. die einzelnen Messzeitpunkte hinweg festzustellen. Zu t_3 zeigt die Kontrollschule (Egbert) signifikant höhere Werte für leistungsmotivationale Integration. Zum vierten und fünften Messzeitpunkt unterscheiden sich die beiden Schulen nicht signifikant voneinander (vgl. Abb. 8.16).

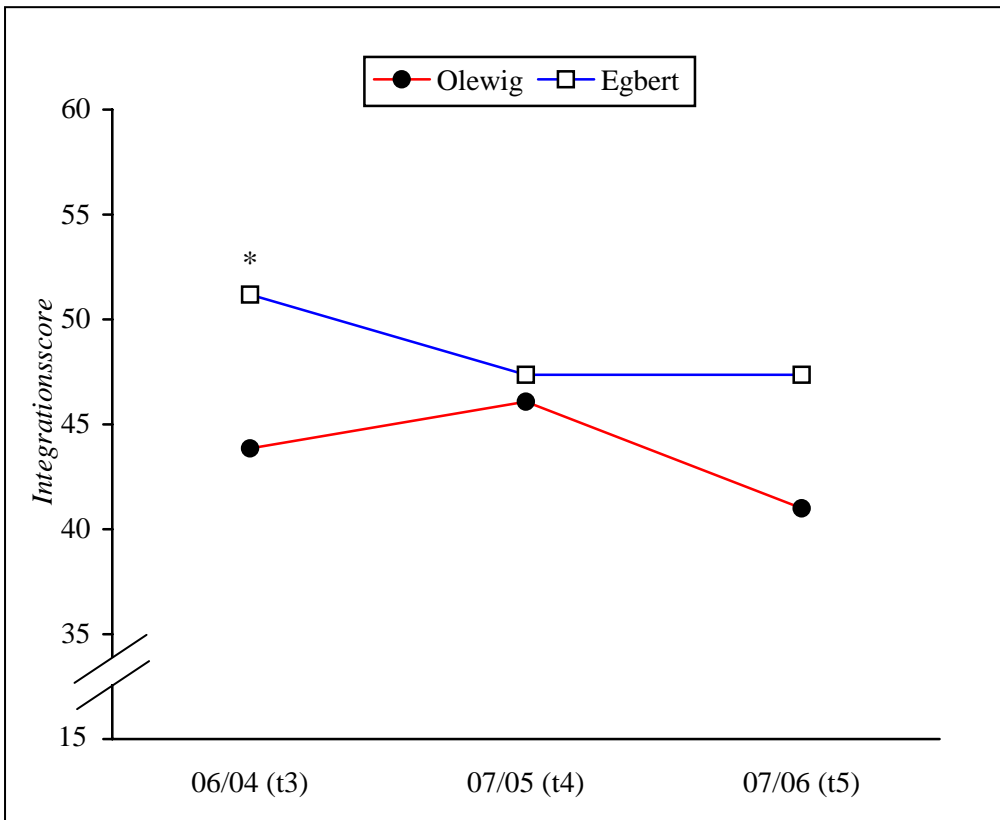


Abbildung 8.16 Scores für leistungsmotivationale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 1 (theoretischer Range: 15 bis 60).

In der Teilstichprobe der Schülerinnen und Schüler, die zu t_3 eher geringe Werte auf der Dimension „*Leistungsmotivationale Integration*“ aufwiesen, findet sich für den dritten und den fünften Messzeitpunkt ein signifikanter Unterschied im Vergleich der beiden Schulen zueinander. In beiden Fällen liegt der Wert der Kontrollschule (Egbert) über dem der Experimentalschule (Olewig). In der Teilstichprobe der hoch integrierten Schülerinnen und Schüler zeigt sich ein bedeutsamer Unterschied zwischen den beiden Schulen zu t_3 . Auch hier ist der Integrationscore der Kontrollschule höher als der Wert der Experimentalschule (vgl. Abb. 8.17).

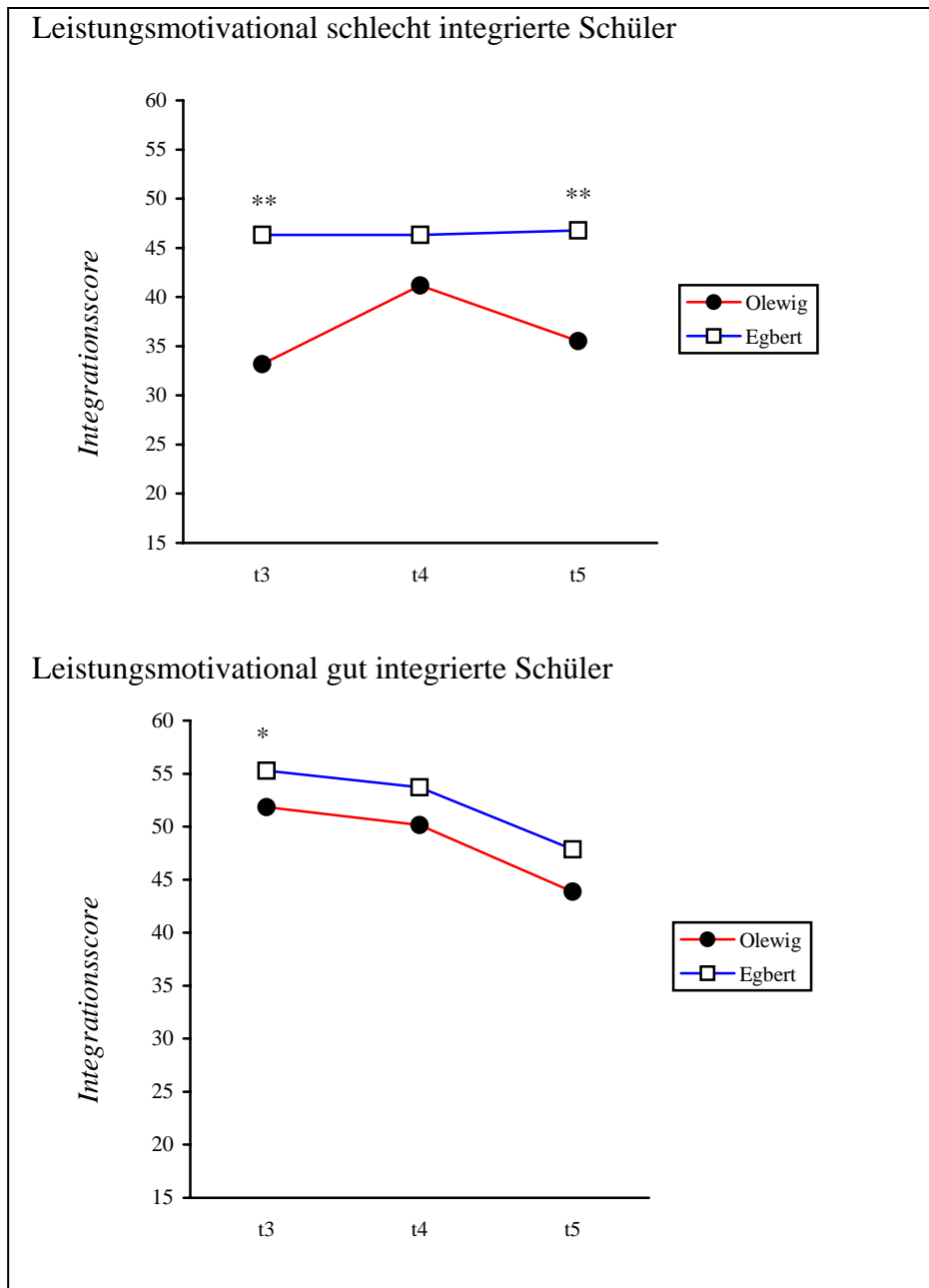


Abbildung 8.17 Scores für leistungsmotivationale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für schlecht integrierte resp. gut integrierte Schüler aus Gruppe 1 (theoretischer Range: 15 bis 60).

Exkurs: Leistungsmessung im Rahmen des „VERA“-Projektes⁵ in Gruppe 1

Projektdarstellung

Im Januar 2007 – also fast etwa zeitgleich mit der Erstellung dieses Abschlussberichtes – wurden erste Ergebnisse des Projektes VERA („Vergleichsarbeiten in der Grundschule“) veröffentlicht, die an dieser Stelle kurz dargestellt werden⁶. Diese Ergebnisse beziehen sich auf Messungen der Leistungsfähigkeit bei jeweils einer Klasse in der Experimental- und Kontrollschule, die zu Beginn unserer Evaluationsstudie im 1. Schuljahr waren und zum Messzeitpunkt der VERA-Erhebung inzwischen das 4. Schuljahr erreicht hatten.

VERA ist ein Gemeinschaftsprojekt von sieben Bundesländern (Berlin, Brandenburg, Bremen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz und Schleswig-Holstein), welches von Ingmar Hosenfeld und Andreas Helmke (beide Universität Koblenz-Landau) in Kooperation mit dem Ministerium für Bildung, Frauen und Jugend, Rheinland-Pfalz (MBFJ) entwickelt und wissenschaftlich durchgeführt wird. Das Projekt zielt darauf ab, vom Jahr 2004 an die Leistungsfähigkeit von Schülerinnen und Schülern in den Fächern Deutsch und Mathematik zu erheben, und zwar in allen beteiligten Bundesländern jeweils zu identischen Zeitpunkten, d.h. zum Beginn der 4. Klassenstufen. (Modifikationen, die ab 2007 geplant sind, werden hier nicht weiter beachtet). Die Aufgabenpools für beide Fächer wurden auf der Grundlage des aktuellen Rahmenplans für die Grundschule erstellt und jeweils im Rahmen von Normierungsstudien mit Zufallsstichproben pro Bundesland geeicht; die Daten dieser Studien sollen die Basis für die Bewertung und relative Einordnung der individuellen Ergebnisse der Schülerinnen und Schüler wie auch der Resultate ganzer Klassen oder Schulen liefern

⁵ Siehe: <http://www.uni-landau.de/vera>

⁶ Die nachfolgende Darstellung beruht ausschließlich auf Informationen, die Herr Kurt Lellinger von der Deutschen Schulschachstiftung dem ZDiag zur Verfügung gestellt hat.

Differenzierung einzelner Fähigkeitsniveaus in vier Leistungsbereichen

Zur Bewertung der Leistungsfähigkeit wurden vorab drei Fähigkeitsniveaus definiert, denen die Schülerinnen und Schüler in den Schulfächern "Deutsch" (mit den Teilbereichen "Leseverständnis" sowie "Sprache und Sprachgebrauch") und "Mathematik" (mit den Teilbereichen "Zahlen und Operationen" und "Größen und Messen") jeweils zugeordnet werden. Diese Fähigkeitsniveaus sollen für die einzelnen Leistungsbereiche im Folgenden kurz dargestellt werden.

Deutsch - Leseverständnis

Um im Bereich Leseverständnis dem *Fähigkeitsniveau 1* zugeordnet zu werden, müssen die Schülerinnen und Schüler über elementare Fähigkeiten im Lesen verfügen. Auf diesem Niveau setzen sie bevorzugt die Technik des re-kodierenden (d.h. innerlich mitsprechenden) Wort-zu-Wort-Lesens ein; die Informationsaufnahme ist eher punktuell orientiert; es werden wörtliche Einzelinformationen verarbeitet und erinnert und um Alltagswissen ergänzt, so dass der Text eher partiell denn als Ganzes verstanden wird. Auf dem *Fähigkeitsniveau 2* verfügen die Kinder über erweitere Fähigkeiten und sind in der Lage, Beziehungen im Text zu erkennen, zu deuten und zu bewerten. In der Interaktion mit dem Text werden Erwartungen der Kinder an den Text mit Gelesenem und Vorwissen verknüpft. Der Text wird in seiner Hauptaussage verstanden. Das reflektierende Lesen erfolgt bezogen auf einzelne Textteile, wobei die Interpretation nicht über ein vordergründiges Verständnis des Textes hinausgeht. Für die Einstufung auf *Fähigkeitsniveau 3* müssen die Kinder Texte in ihrer Gesamtheit und Komplexität verstehen. Sie können überfliegend lesen und verknüpfen unterschiedliche Informationen mit speziellem Vor-, Sach- und Weltwissen und bewerten die zentralen Aussagen. Das reflektierende Lesen ist weiter ausgebildet, so dass Schlussfolgerungen, die über den Text hinausgehen, gezogen werden können. Zudem sind die Kinder auf diesem Niveau in der Lage, einzelne Textstellen gezielt zu verarbeiten und zu interpretieren, um zu einem Gesamtverständnis des Textes zu gelangen.

Deutsch: Sprache und Sprachgebrauch

Die drei Fähigkeitsniveaus in diesem Bereich sind durch qualitativ unterschiedliche Strategien im Umgang mit sprachlichen Phänomenen durch das Kind gekennzeichnet. Die Fähigkeiten, die das jeweils voraus laufende Niveau charakterisierten, sind im nächst höheren Niveau eingeschlossen. Kinder auf *Fähigkeitsniveau 1* können eindeutige (grammatische) Fehler spontan korrigieren, Wortbildungen und -formen nach Muster leisten sowie ein Wort oder einen Ausdruck in den Kontext angemessen einsetzen. Sie arbeiten weitgehend spontan aus dem Sprachgefühl heraus, allerdings ohne große Überlegungen und Bewusstheit. Kinder, die auf *Fähigkeitsniveau 2* eingestuft werden, arbeiten mit entwickelter Aufmerksamkeit für Sprachphänomene nach "inneren Regeln" über Sprache. Die Sprachbewusstheit lässt sich etwa als "analytisch geleitetes Sprachgefühl" beschreiben, das selbst erarbeitetes, implizites (d.h. nicht formuliertes) Wissen über Sprache, welches regelhaft "gelebt" wird, enthält. Für eine Einstufung auf

Fähigkeitsniveau 3 müssen auch im Rahmen offener Aufgaben komplexe grammatische und stilistische Phänomene erfolgreich bearbeitet werden können. Schülerinnen und Schüler auf diesem Niveau handeln überlegt und wenden das im Unterricht erworbene Wissen bewusst und zielgerichtet an. So ist beispielsweise die Untersuchung auf die Funktion eines sprachlichen Mittels oder einer Form möglich. Die Sprachbewusstheit enthält explizit formulierbares Wissen über Sprache und Sprachphänomene.

Mathematik: Zahlen und Operationen

Für diesen Bereich gelten als Kriterien für die Einstufung auf dem *Fähigkeitsniveau 1*, dass Zahlen aus gebräuchlichen Darstellungen abgelesen werden können und die Zuordnung von Zahlen zu Zahlendarstellungen und umgekehrt beherrscht wird; die Grundaufgaben des Kopfrechnens (Einspluseins, Einmaleins, Zahlzerlegungen) und die sichere Ableitung von Umkehrungen werden für Addition und Subtraktion mit überschaubarem Zahlenmaterial, bei Multiplikation und Division bei einfachen Aufgaben beherrscht; das schriftliche Verfahren der Addition ohne Übertrag wird verstanden und korrekt angewendet; Lösungen können für die Addition durch Überschlagsrechnung kontrolliert werden. Einsicht in das Stellenwertsystem liegt vor und kann bei einfachen Additionsaufgaben mit Überschreiten und bei Subtraktionsaufgaben ohne Überschreiten des Tausenders angewendet werden. Auf dem *Fähigkeitsniveau 2* können Zahlen auch aus weniger gebräuchlichen Darstellungen abgelesen werden, und die Zuordnung von Zahlen zu Zahlendarstellungen und umgekehrt wird beherrscht. Um Zielzahlen zu erhalten, können Zahlen in verschiedenen Darstellungen abgelesen und zueinander in Beziehung gesetzt werden. Die Grundaufgaben des Kopfrechnens (Einspluseins, Einmaleins, Zahlzerlegungen) und die sichere Ableitung von Umkehrungen werden für alle Grundrechenarten auch bei komplexeren Aufgaben (Einsetzen von fehlenden Rechenoperationen in Gleichungen, weniger gebräuchliche Formate) mit einfachen Texten beherrscht. Das schriftliche Verfahren der Addition wird verstanden und auch mit Überträgen und in weniger vertrauten Formaten korrekt angewendet; die Anwendung von Überschlagsrechnungen und Umkehroperationen für die Addition gelingt. Es liegt Einsicht in das Stellenwertsystem liegt vor, und es kann bei einfachen Additions- und Subtraktionsaufgaben auch bei Überschreitung des Tausenders sowie beim Runden angewandt werden. Rechenstrategien werden verstanden und bei geeigneten Aufgaben angewendet. Kontextbezogene Aufgaben werden auch bei sprachlich komplexeren oder längeren Texten gelöst, und dabei kann die Beziehung zwischen der Sache und den einzelnen Lösungsschritten beschrieben werden. Für die Zuordnungen der individuellen Leistungen zu dem *Fähigkeitsniveau 3* sind folgende Kriterien definiert: Die Grundaufgaben des Kopfrechnens (z.B. Einmaleins, Zahlzerlegungen) und die sichere Ableitung von Umkehrungen werden für alle Grundrechenarten auch bei komplexeren Aufgaben (Einsetzen von fehlenden Rechenoperationen in Gleichungen, weniger gebräuchliche Formate) mit sprachlich anspruchsvolleren oder längeren Texten beherrscht. Dabei können Zahlen und Operationen flexibel kombiniert werden. Die schriftlichen Verfahren der Addition und Subtraktion werden verstanden und auch mit schwierigen Überträgen in weniger gebräuchlichen Formen korrekt angewendet. Die Multiplikation wird beherrscht und bei geeigneten Aufgaben korrekt angewendet. Die Anwendung von Überschlagsrechnungen gelingt für alle Grundrechenarten. Einsicht in das Stellenwertsystem liegt vor, und Veränderungen können durch geeignete arithmetische Operationen flexibel vorgenommen werden. Eine kombinatorische

Problemstellung kann vollständig modelliert werden. Die vier Grundrechenarten können flexibel kombiniert unter Einsatz mathematischer Kenntnisse (z.B. Rechengesetze, Teilbarkeitsregeln) werden.

Mathematik: Größen und Messen

Es liegen uns keine Informationen darüber vor, wie dieser Bereich erfasst wurde oder welche Anforderungen mit den einzelnen Fähigkeitsniveaus verknüpft sind.

VERA-Ergebnisse für Gruppe 1 der Evaluationsstudie

In der folgenden Tabelle sind die Ergebnisse, wie sie im Rahmen von VERA ermittelt wurden, für die beiden vierten Klassen in Olewig und Egbert, die auch an der Evaluationsstudie teilgenommen haben, aufgeführt. Mitgeteilt wird die prozentuale Verteilung der Schülerinnen und Schüler jeder der beiden Schulen auf die oben genannten Fähigkeitsniveaus in den vier Leistungsbereichen. Mitgeteilt werden auch die entsprechenden Zahlen für die Schülerinnen und Schüler, die landesweit in Rheinland-Pfalz an dem Projekt teilgenommen haben und deren Ergebnisse die Hintergrundfolie für einen Vergleich mit dem „Landesdurchschnitt“ bilden.

Tabelle 8.1. Verteilung der Schülerinnen und Schüler beider Schulen auf drei Fähigkeitsniveaus in vier Leistungsbereichen sowie Vergleichszahlen aus der Gesamterhebung für Rheinland-Pfalz (alle Angaben in Prozent)*

Fähigkeitsniveau	„Experimentalschule“ Olewig	„Kontrollschule“ Egbert	Rheinland-Pfalz
Deutsch: Leseverständnis			
1	10	29	39
2	31	48	35
3	57	18	20
nicht bestimmbar*	-	3	6
Deutsch: Sprache und Sprachgebrauch			
1	0	18	16
2	15	44	49
3	84	37	32
nicht bestimmbar*	-	-	3
Mathematik: Zahlen und Operationen			
1	21	17	23
2	21	39	46
3	57	42	38
nicht bestimmbar*	-	-	-
Mathematik: Größen und Messen			
1	10	10	25
2	36	67	50
3	52	21	24
nicht bestimmbar*	-	-	1

Eine genauere Inspektion des Zahlenmaterials zeigt, dass im Bereich „*Deutsch: Leseverständnis*“ die Leistungen der derzeit 4. Klasse der Grundschule Olewig bedeutsam über dem Landesdurchschnitt liegen; die Leistungen der Schülerinnen und Schüler der 4. Klasse der Grundschule Egbert unterscheiden sich dagegen nicht bedeutsam vom Landesschnitt. Ob die Leistungen der beiden Klassen aus Olewig bzw. Egbert sich im Vergleich zueinander bedeutsam unterscheiden, kann aufgrund der vorliegenden Informationen nicht bestimmt werden. Auch in dem Leistungsbereich „*Deutsch: Sprache und Sprachgebrauch*“ liegen die Leistungen der

Viertklässler(innen) der Grundschule Olewig bedeutsam über dem landesweiten Durchschnitt. Für die Grundschule Egbert gilt dies erneut nicht:

Für den Leistungsbereich „*Mathematik: Zahlen und Operationen*“ ergibt sich das gleiche Bild: Auch hier unterscheidet sich nur die Leistung der Schülerinnen und Schüler der vierten Klasse aus Olewig bedeutsam vom landesweiten Durchschnitt, nicht jedoch die Leistung der vierten Klasse in Egbert. Schließlich gilt auch für den Bereich „*Mathematik: Größen und Messen*“ das oben Gesagte: Auch hier liegt die Leistung der Schülerinnen und Schüler aus Olewig bedeutsam über dem Landesdurchschnitt, während sich die Werte für die Schülerinnen und Schüler aus Egbert nicht bedeutsam vom landesweiten Mittelwert unterscheiden.

Kommentar

Summarisch lässt sich festhalten, dass eine deutliche Überlegenheit der Schülerinnen und Schüler der vierten Grundschulklasse Olewig in allen hier betrachteten Leistungsbereichen zu erkennen ist: Ihre Leistungen liegen deutlich über dem landesweiten Durchschnitt, während die Leistungen der Schülerinnen und Schüler aus der Egbert-Schule sich nicht vom landesweiten Durchschnitt unterscheiden. Die hier interessierende Frage ist nun, ob und inwieweit dieses Leistungsniveau der Olewiger Schülerinnen und Schüler mit dem Schachunterricht, den sie seit ihrer ersten Klasse in der Schule erhalten haben, zusammenhängt oder gar durch diesen „erzeugt“ worden ist. Vor Beantwortung dieser Frage gilt es allerdings zu prüfen, inwieweit die VERA-Daten es erlauben, dass man sie als mögliche Belege für die förderlichen Wirkungen des Schachunterrichts in der Grundschule Olewig heranzieht resp. inwieweit mögliche Einschränkungen der Aussagekraft dieser Befunde angenommen werden müssen. Dies soll im folgenden kurz skizziert werden.

- (1) Die Aufgaben in dem VERA-Projekt werden von den entsprechenden Klassenlehrerinnen und Klassenlehrern selbst vorgegeben und auch selbst ausgewertet. Daher liegt strenggenommen keine *standardisierte* und keine *anonymisierte* Datenerhebung und -auswertung vor, wie sie in einer Evaluationsstudie realisiert werden müssen.
- (2) Es kann nicht ganz ausgeschlossen werden, dass die Lehrerinnen und Lehrer die entsprechenden Testaufgaben vor dem offiziellen Erhebungszeitpunkt aus

- dem Internet heruntergeladen und mit ihren Schülern ähnliche Übungsaufgaben bearbeitet oder sie entsprechend unterrichtet haben.
- (3) Da das VERA-Projekt nur Schülerinnen und Schüler der vierten Klassen einbezieht, kann der Vergleich nicht ausgeweitet werden auf (jüngere) Kinder aus anderen Klassen, die bislang keinen Schachunterricht erhielten. Die überdurchschnittlich guten Leistungen der jetzigen Viertklässler der Grundschule Olewig könnten nur dann mit dem Schachunterricht in Verbindung gebracht (oder auf diesen sogar zurückgeführt) werden, wenn die jüngeren Klassen, d.h. Klasse eins, zwei und drei der Grundschule Olewig *nicht* bedeutsam über dem Landesdurchschnitt lägen. Sollten diese im (fiktiven) Vergleich mit dem Landesdurchschnitt jedoch ebenso überdurchschnittliche Leistungen zeigen wie die jetzigen Viertklässler, dann wäre nicht der Schulschachunterricht, sondern wären eher schulspezifische Faktoren für die überdurchschnittlichen Leistungen ursächlich heranzuziehen (z.B. Qualität des Unterrichts, Besonderheiten bei der VERA-Datenerhebung). Da ein solcher Vergleich der jetzigen Viertklässler mit unteren Klassen in Olewig ist jedoch nicht möglich ist, lässt sich deren überdurchschnittliche Leistungsfähigkeit zwar *aus Plausibilitätsgründen* auf den Schachunterricht zurückführen, nicht aber wirklich belegen.
- (4) In den vorliegenden Daten ist jeweils nur angegeben, ob sich die Werte der einzelnen Klassen signifikant vom Landesdurchschnitt unterscheiden. Es ist jedoch nicht bekannt, ob sich die beiden vierten Klassen aus Olewig und Egbert auch im direkten Vergleich bedeutsam voneinander unterscheiden. Feststeht nur, dass sich die Schülerinnen und Schüler aus Olewig zwar in allen Leistungsbereichen bedeutsam vom Landesdurchschnitt unterscheiden und dass dies nicht auf die Kinder aus Egbert zutrifft; fest steht auch, dass das durchschnittliche Leistungsniveau der Kinder, die die Grundschule in Olewig besuchen, über dem der Kinder aus der Egbert-Schule liegt; es liegen aber keine Daten dazu vor, ob dieser Leistungsunterschied zwischen den Schulen auch statistisch bedeutsam, d.h. nicht auf Zufall zurückzuführen ist.
- (5) Bei der Beschreibung der Ziele von VERA wird ausdrücklich darauf verwiesen, dass ein Ranking von Schulen nicht vorgesehen ist, sondern dass eher innerschulische Vergleiche angestrebt werden sollen. Ein Vergleich zweier

Klassen aus unterschiedlichen Schulen ist somit zwar deskriptiv möglich, wird aber grundsätzlich weder empfohlen noch kann er als zuverlässig gelten.

Als Fazit aus dem Gesagten lässt sich festhalten, dass die Daten aus der VERA-Studie für die im Kontext der Schachevaluation zentrale Fragestellung nur beschränkten Aussagewert besitzen können und dass sie keine gesicherten Schlussfolgerungen ob der möglichen förderlichen Auswirkungen des Schachunterrichts in der Grundschule Olewig erlauben.

8.2 Befunde für Gruppe 2

Die Schülerinnen und Schüler der Gruppe 2 besuchten zu Beginn der Evaluationsstudie das zweite und zum fünften Messzeitpunkt das vierte Schuljahr. Die Schülerinnen und Schüler der Experimentalschule (Olewig) aus dieser Gruppe erhielten bis zum fünften Erhebungszeitpunkt ebenfalls drei Jahre lang Schachunterricht.

8.2.1 Niveau der allgemeinen Intelligenz

Im Vergleich der beiden Schulen zueinander zeigen die Kinder mit Schachunterricht (Olewig) zu t_2 und t_3 signifikant bessere Testleistungen in der *Grundintelligenztestskala 1 (CFT 1)* als die der Kontrollschule (Egbert); auch zeigen die Leistungen der Olewiger Kinder über die ersten Messzeitpunkte hinweg einen signifikanten Anstieg, während dieser bei den Egbert-Kindern nicht statistisch bedeutsam ist. Anfänglich wie auch zu den beiden letzten Messzeitpunkten sind keine bedeutsamen Unterschiede mehr zwischen den Schulen feststellbar (vgl. Abb. 8.18).

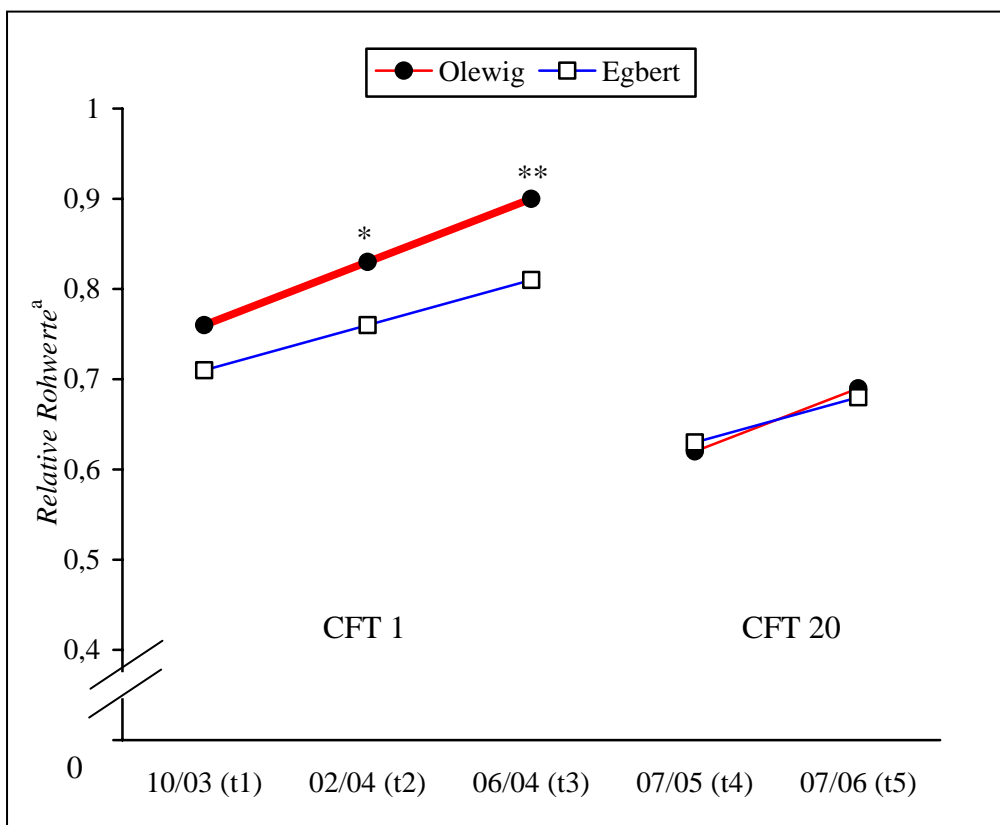


Abbildung 8.18 Gesamtscore der allgemeinen Intelligenz im CFT1 ($t_1 - t_3$) und CFT20 ($t_4 - t_5$) über die fünf Messzeitpunkte für Gruppe 2.

^aDa zum Testzeitpunkt t_4 und t_5 eine andere Testversion mit einem anderen Wertebereich verwendet wurde, erfolgt die Darstellung aus Gründen der Vergleichbarkeit mit relativen Rohwerten.

Die getrennte Betrachtung der leistungsstarken und leistungsschwachen Kinder zeigt, dass letztere in der Experimentalschule (Olewig) zu t_3 signifikant höhere Werte erzielen und dass sich hier die Leistungen von t_1 nach t_2 und von t_2 nach t_3 bedeutsam steigern, während in der Kontrollschule nur der Zuwachs von t_1 nach t_2

signifikant ist. Die signifikant niedrigeren Leistungen zum 4. und 5. Messzeitpunkt sind wie in Gruppe 1 mit dem Wechsel der Testversion begründet; signifikante Unterschiede zwischen den Schulen zeigen sich nicht (vgl. Abb. 8.19).

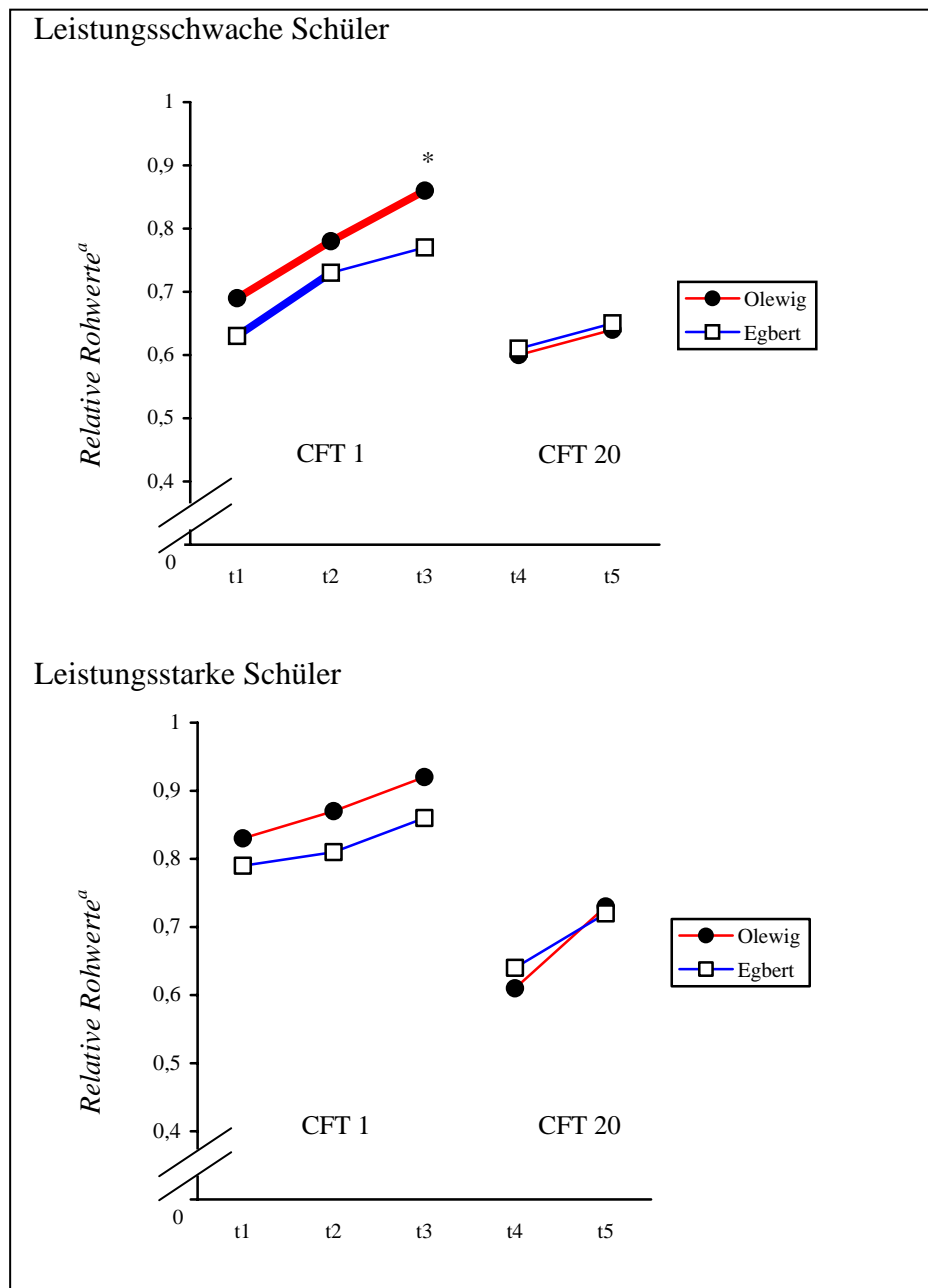


Abbildung 8.19 Gesamtscore der allgemeinen Intelligenz im CFT1 (t₁ - t₃) und CFT20 (t₄ - t₅) über die fünf Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 2.

^aDa zum Testzeitpunkt t₄ und t₅ eine andere Testversion mit einem anderen Wertebereich verwendet wurde, erfolgt die Darstellung aus Gründen der Vergleichbarkeit mit relativen Rohwerten.

Für einzelne Untertests, nämlich v.a. für „*Substitution*“ und „*Labyrinthe*“ lassen sich signifikante Differenzen zugunsten der Olewiger Kinder feststellen (Abb. 8.20), die insbesondere auf einen anfänglich sehr signifikanten Anstieg der Leistungen bei den als „leistungsschwach“ eingestuften Kindern zurück gehen; entsprechende Leistungszuwächse über die Zeit zeigen sich auch für Egbert, sie liegen jedoch im Niveau unter denen in Olewig (Abb. 8.21). In anderen Untertests (v.a. Klassifikation und „Ähnlichkeiten“) nehmen die Leistungen in beiden Schulen gleichermaßen zu, was auch hier wieder auf einen relativen Anstieg v.a. bei den leistungsschwachen Schülerinnen und Schüler in *beiden* Schulen zurückzuführen ist.) einen signifikanten Leistungszuwachs von t_4 nach t_5 , während sich für die Teilstichprobe der leistungsstarker Schülerinnen und Schüler kaum signifikante Veränderungen zeigen.

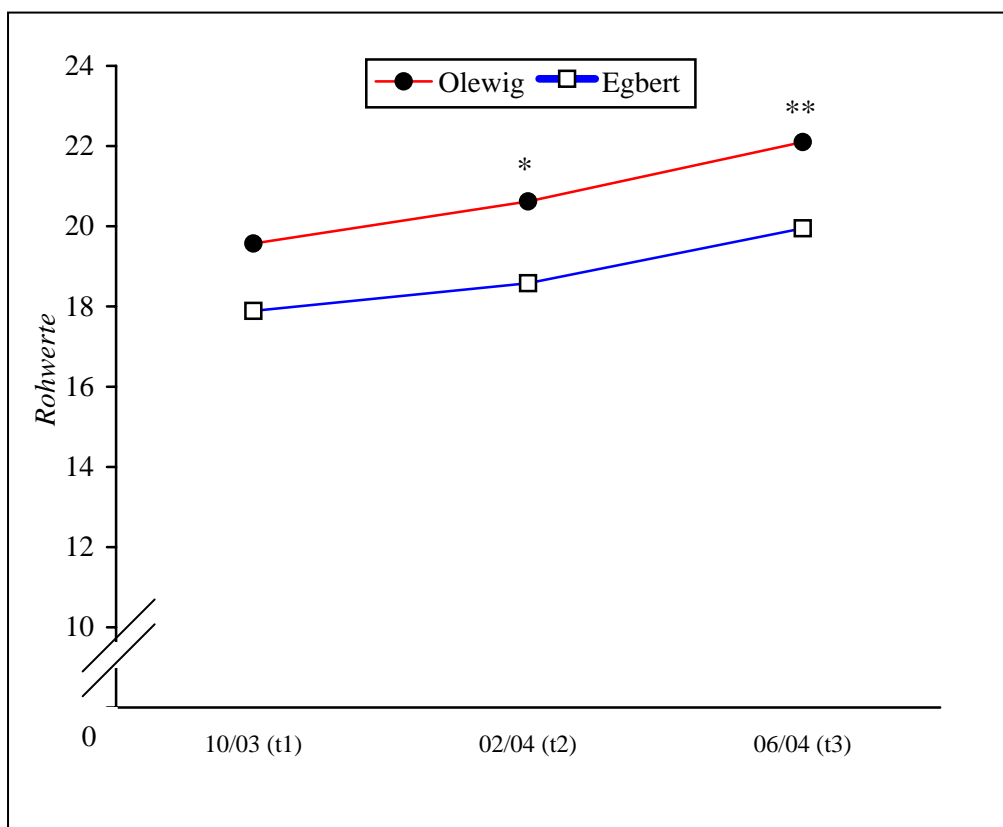


Abbildung 8.20 Summe der Rohwerte in den Untertests „*Substitution*“ und „*Labyrinthe*“ des CFT1 über die drei Messzeitpunkte für Gruppe 2 (theoretischer Range: 0 bis 24). Mit dem zu t_4 und t_5 verwendeten CFT 20 werden keine entsprechenden Werte erfasst.

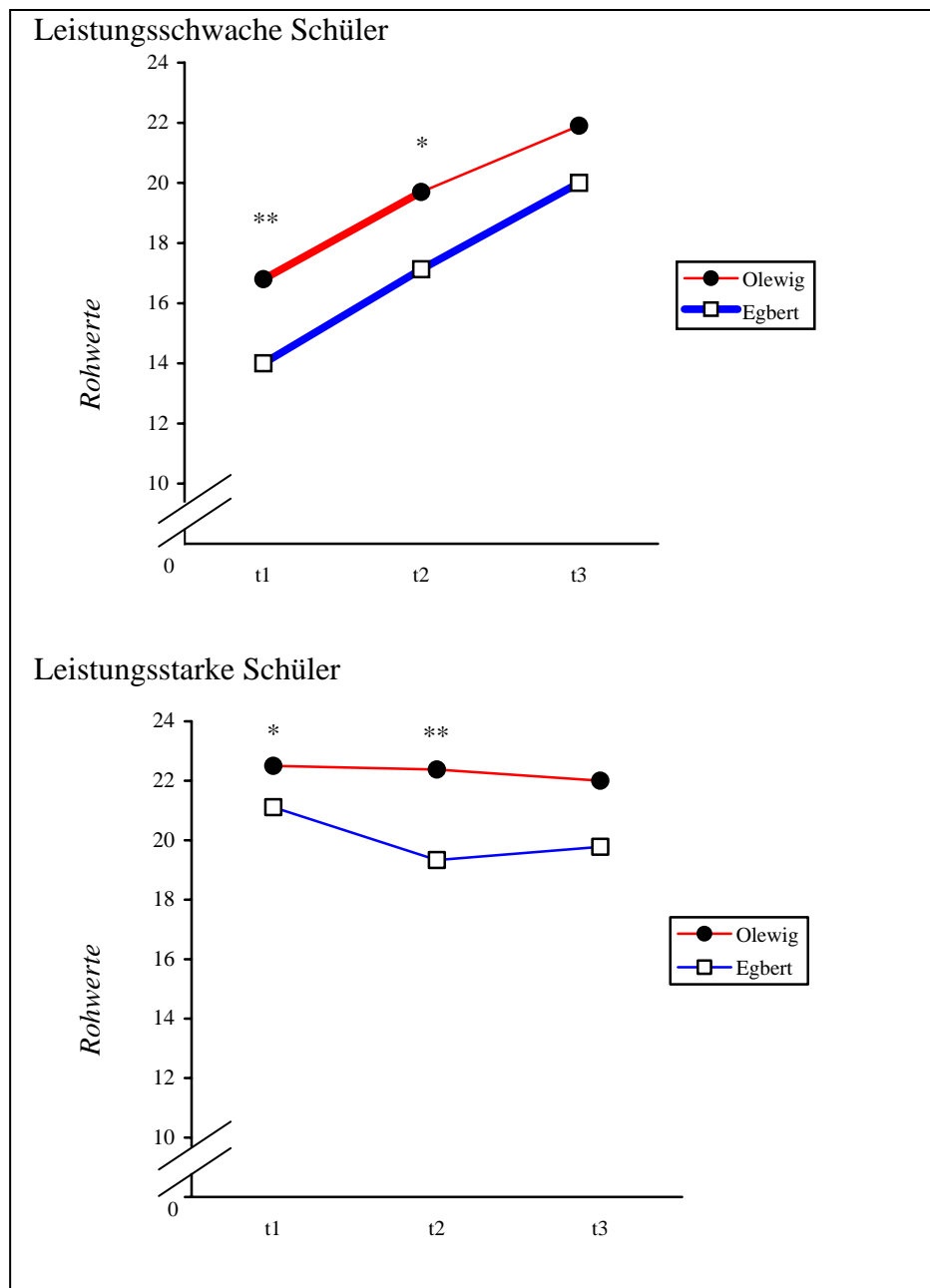


Abbildung 8.21 Summe der Rohwerte in den Untertests „Substitution“ und „Labyrinth“ des CFT1 über die drei Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 2 (theoretischer Range: 0 bis 24). Mit dem zu t_4 und t_5 verwendeten CFT 20 werden keine entsprechenden Werte erfasst.

Im *Untertest „Matrizen“* (als Indikator Regelhaftigkeiten bei figuralem Material zu erkennen) zeigt sich ein deutlicher Leistungsanstieg der Olewiger Kinder zwischen den beiden letzten Erhebungszeitpunkten, und zwar ausschließlich bei den leistungsschwachen Kindern; für die leistungsstarken Kinder zeigt sich in Egbert ein signifikanter *Leistungsabfall* (vgl. Abb. 8.22), der in Olewig nicht zu beobachten ist.

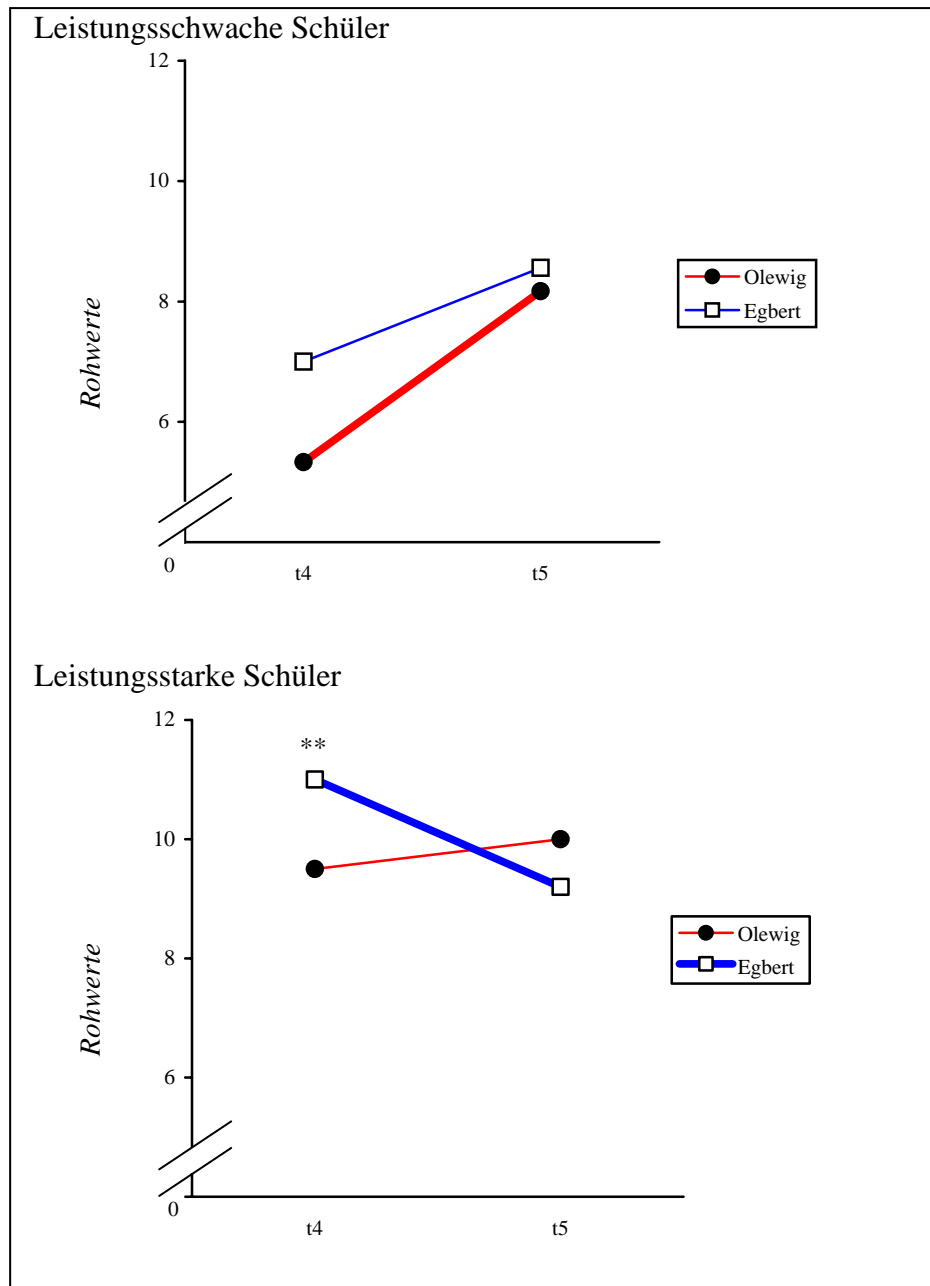


Abbildung 8.22 Absolute Rohwerte im Untertest „Matrizen“ des CFT20 über die beiden Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 2 (theoretischer Range: 0 bis 12). Mit dem zu den ersten drei Messzeitpunkten verwendeten CFT 1 werden keine entsprechenden Werte erfasst.

8.2.2 Konzentrationsfähigkeit

Im *Differentiellen Leistungstest (DL-KE / DL-KG)* zeigen *alle* Schülerinnen und Schüler über den Verlauf des ersten Studienjahres einen signifikanten Anstieg ihrer Konzentrationsfähigkeit, der sich jedoch danach noch bei den Schülerinnen und Schülern mit Schachunterricht, nicht mehr aber bei jene ohne Schachunterricht festsetzt, sodass die Kinder aus Olewig zu t_4 signifikant bessere Leistungen erreichen als die der Kontrollschule (Egbert); allerdings lässt sich dieser Unterschied zu t_5 nicht mehr nachweisen (vgl. Abb. 8.23).

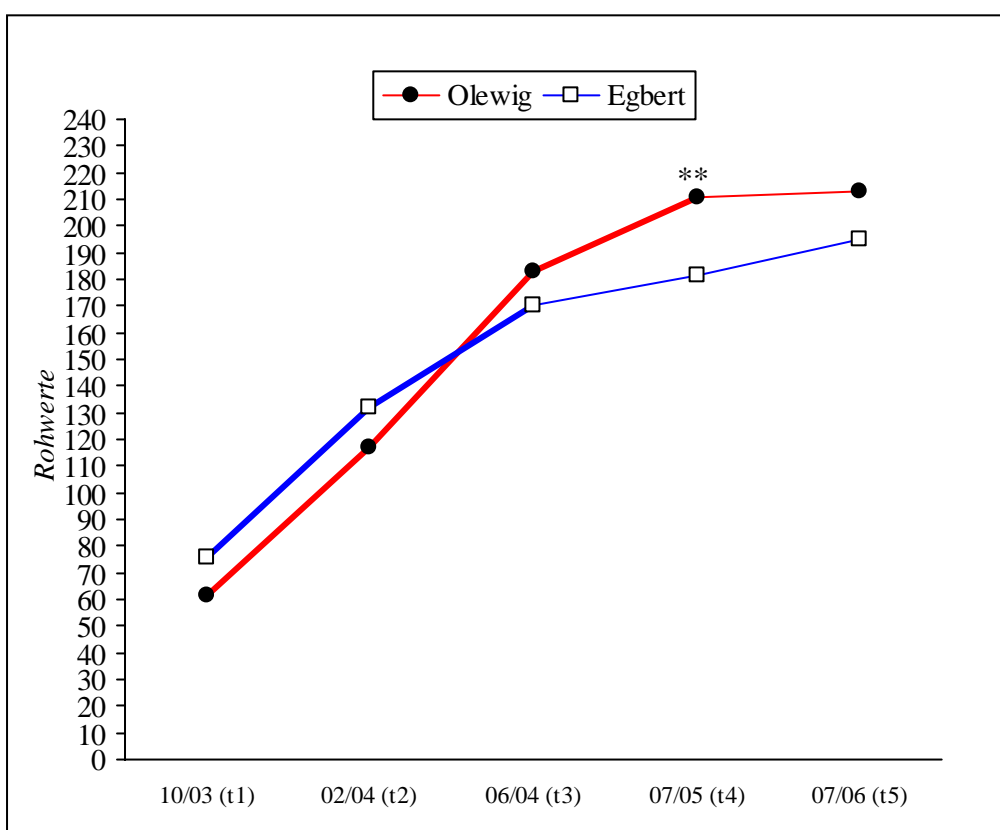


Abbildung 8.23 Absolute Rohwerte im DL-KG über die fünf Messzeitpunkte (t_1 - t_5) für Gruppe 2 (theoretischer Range: 0 bis 240).

Ein fast ähnliches Bild ergibt sich, wenn man den Verlauf für die (anfänglich) leistungsschwachen Kinder betrachtet: Diejenigen, die Schachunterricht erhalten (Olewig), zeigen einen Anstieg ihrer Konzentrationsfähigkeit von t_1 nach t_2 , von t_2 nach t_3 und von t_3 nach t_4 , während dieser Anstieg für die entsprechend leistungsschwachen Kinder der Kontrollschule nur anfänglich, nicht mehr aber von t_3 nach t_4 nachweisbar ist und hier die Konzentrationsleistungen in der

Experimentalschule (Olewig) signifikant über denen der Kontrollschule (Egbert) liegen. Bei den leistungsstarken Kindern zeigt sich in beiden Schulen ein signifikanter Anstieg der Leistung von t_1 nach t_2 und von t_2 nach t_3 , während auch bei diesen nur die Experimentalschule einen weiteren substantiellen Leistungszuwachs von t_3 nach t_4 verzeichnen kann und die Konzentrationsfähigkeit der leistungsstarken Kinder der Experimentalschule über der bei den Kindern der Kontrollschule liegt (vgl. Abb. 8.24).

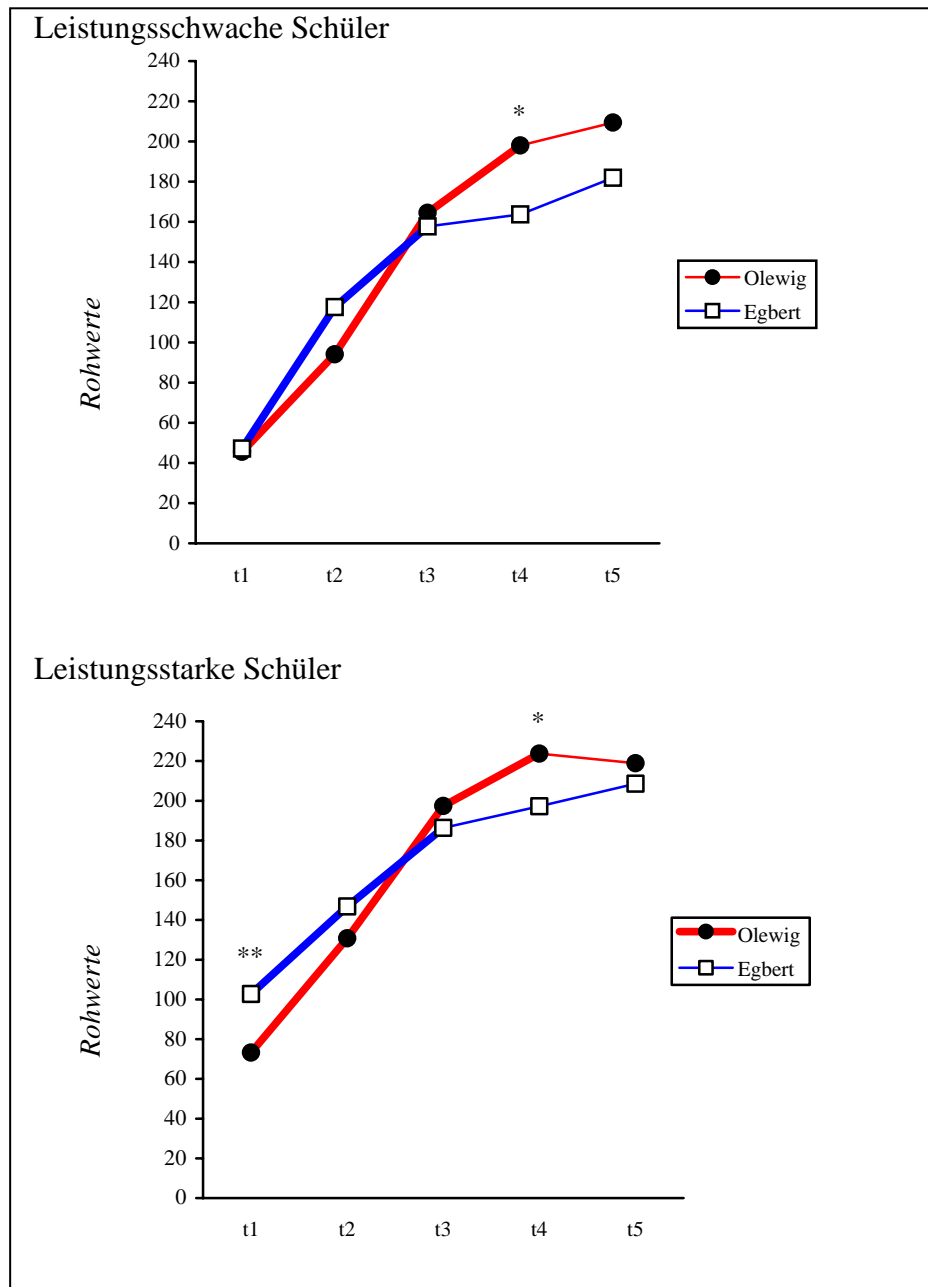


Abbildung 8.24 Absolute Rohwerte im DL-KG über die fünf Messzeitpunkte (t_1 - t_5) für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 2 (theoretischer Range: 0 bis 240).

In der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* erzielen bezogen auf das Maß „Anzahl abgeschriebener Silben“ die Kinder in beiden Schulen einen signifikanten Leistungszuwachs zwischen t_3 und t_4 , wobei allerdings die Schülerinnen und Schüler mit Schachunterricht jeweils höhere Werte aufweisen als die Kinder ohne Schachunterricht (vgl. Abb. 8.25). Indes erbringt die Analyse auch einen bedeutsamen Anstieg der Fehlerzahl in der Experimentalschule von t_3 nach t_4 , so dass zu t_4 die Anzahl der Fehler im Abschreibtext bei den Kindern mit Schachunterricht signifikant höher ist als bei denen der Kontrollschule (vgl. Abb. 8.26). Mit Blick auf die rezeptive Aufmerksamkeit zeigen sich keine bedeutsamen Veränderungen der Leistung zwischen den einzelnen Messzeitpunkten. Zum dritten Messzeitpunkt liegt die rezeptive Aufmerksamkeit in der Experimentalschule (Olewig) signifikant über der der Kontrollschule (Egbert). Zu den späteren Erhebungszeitpunkten ist der Unterschied allerdings nicht mehr statistisch bedeutsam (vgl. Abb. 8.27).

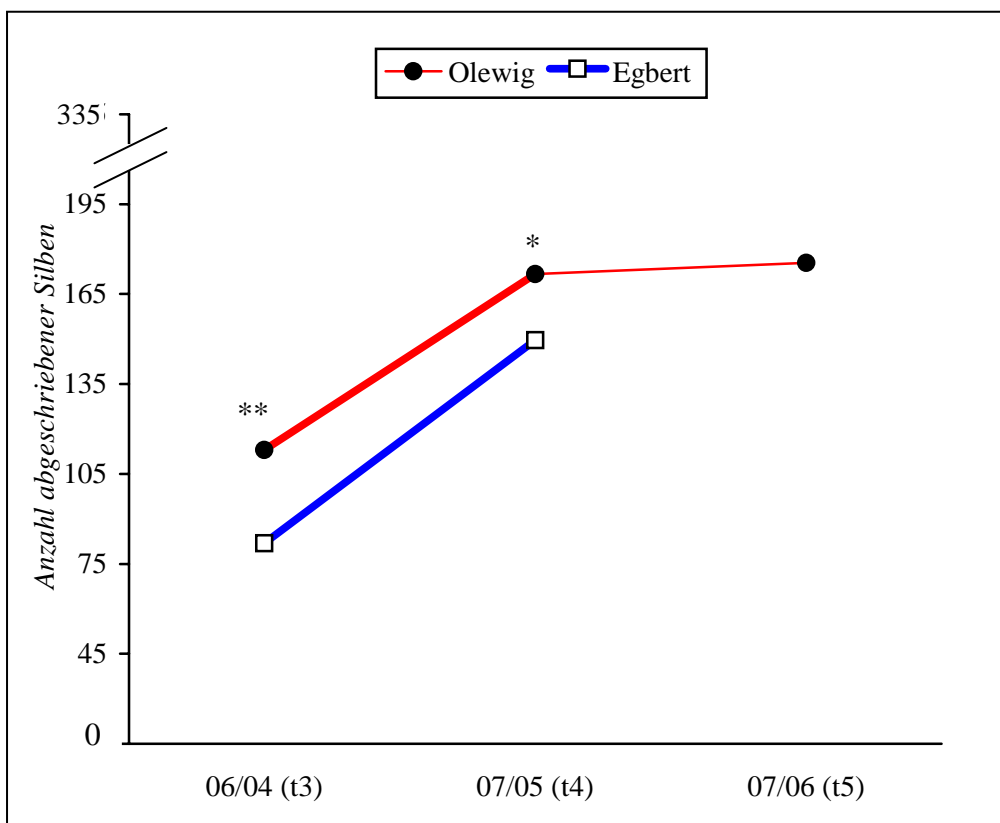


Abbildung 8.25 Anzahl abgeschriebener Silben im TPK über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 2 (theoretischer Range: 0 bis 335). Für den fünften Messzeitpunkt liegen nur die Daten für die Schülerinnen und Schüler der Grundschule Olewig vor.

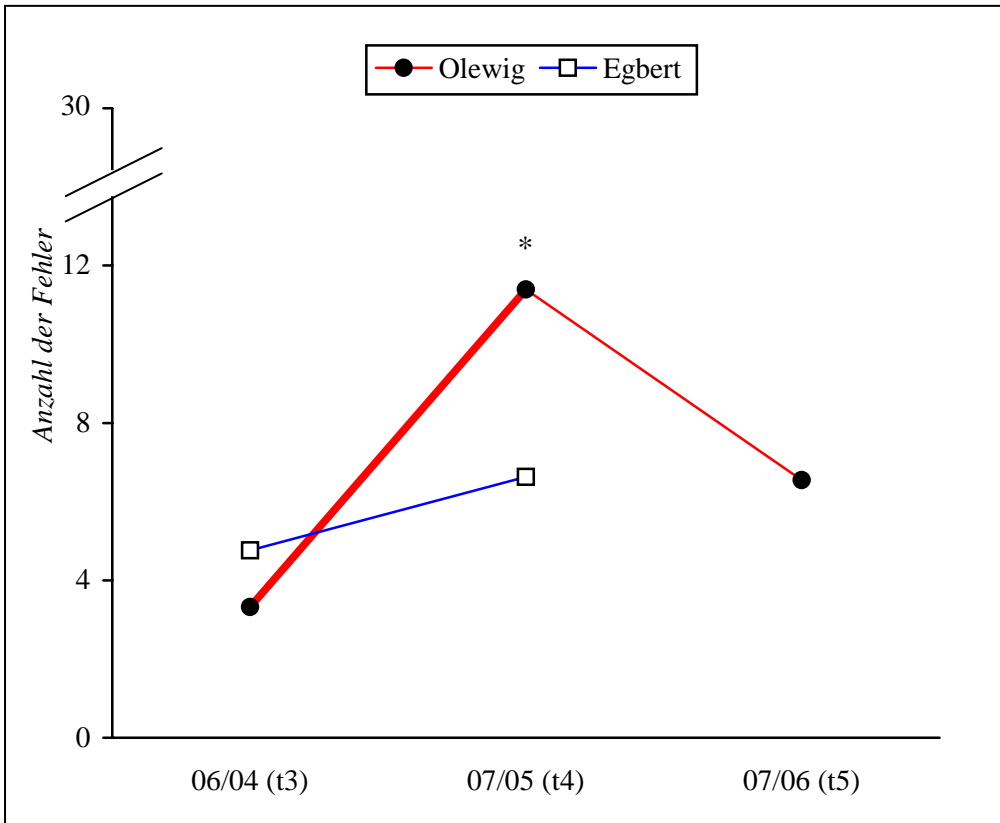


Abbildung 8.26 Anzahl der Fehler im Abschreibtext des TPK über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 2. Für den fünften Messzeitpunkt liegen nur die Daten für die Schülerinnen und Schüler der Grundschule Olewig vor.

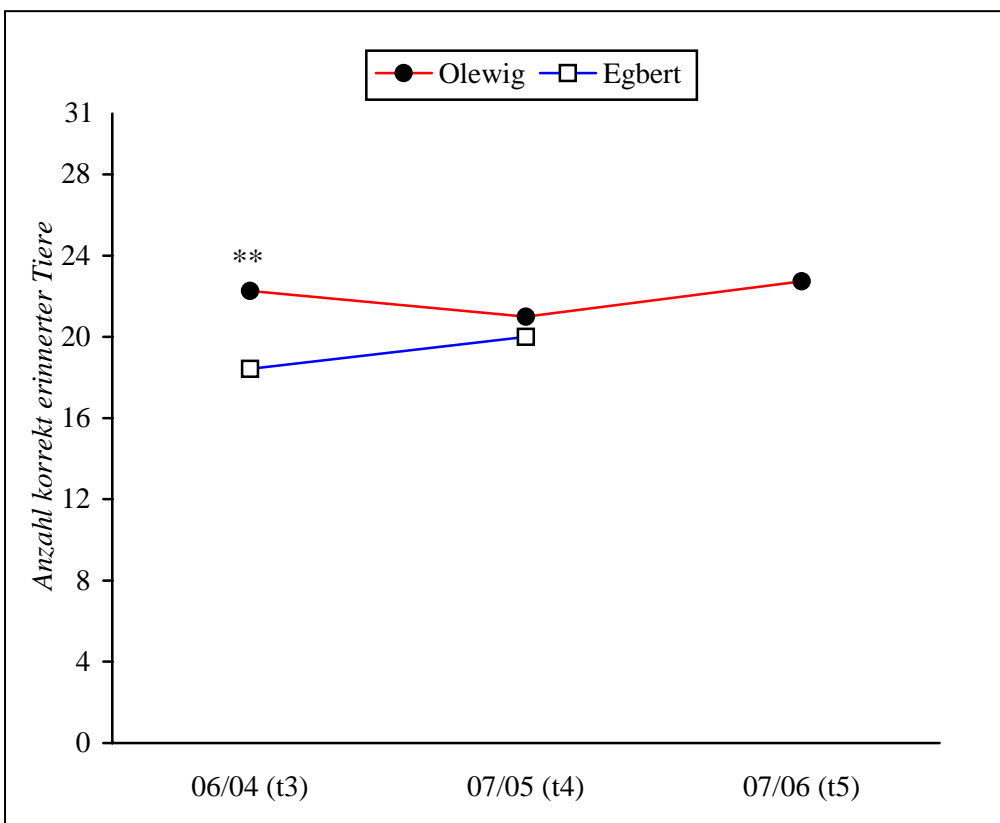


Abbildung 8.27 Anzahl korrekt erinnelter Tiere im TPK über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 2 (theoretischer Range: 0 bis 31). Für den fünften Messzeitpunkt liegen nur die Daten für die Schülerinnen und Schüler der Grundschule Olewig vor.

8.2.3 Prüfung von Rechen- und Rechtschreibleistungen

Ein signifikanter Leistungsanstieg in der Gesamtzahl der bearbeiteten Rechenaufgaben in der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* findet sich nur von t_3 nach t_4 bei den Schülerinnen und Schülern ohne Schachunterricht (Egbert), wohl aber liegen zu beiden Messzeitpunkten die Leistungen in der Experimentalschule (Olewig) höher (vgl. Abb. 8.28).

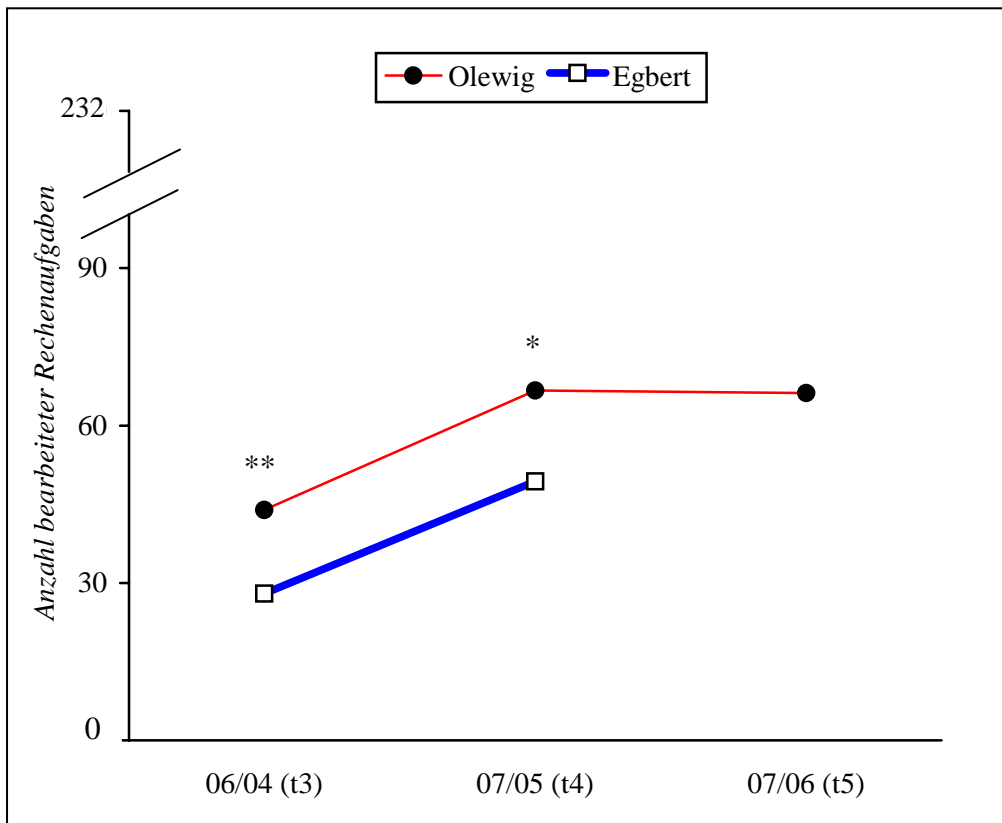


Abbildung 8.28 Anzahl bearbeiteter Rechenaufgaben im TPK über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 2 (theoretischer Range: 0 bis 232). Für den fünften Messzeitpunkt liegen nur die Daten für die Schülerinnen und Schüler der Grundschule Olewig vor.

Eine bedeutsame Veränderung in der Zahl der Fehler über die Zeit lässt sich für keine der beiden Schulen feststellen. Auch der Vergleich der beiden Gruppen zueinander ergibt zu keinem der Messzeitpunkte bedeutsame Unterschiede.

Für die Rechtschreibleistung in der *Hamburger Schreibprobe (HSP)* ist für die Experimentalschule (Olewig) ein signifikanter Leistungszuwachs von t_1' nach t_3 und von t_3 nach t_4 feststellbar. In der Kontrollschule findet sich nur von t_3 nach t_4 ein bedeutsamer Anstieg. Im Vergleich der beiden Schulen zueinander zeigen sich zu keinem der beiden Zeitpunkte bedeutsamen Unterschiede (vgl. Abb. 8.29).

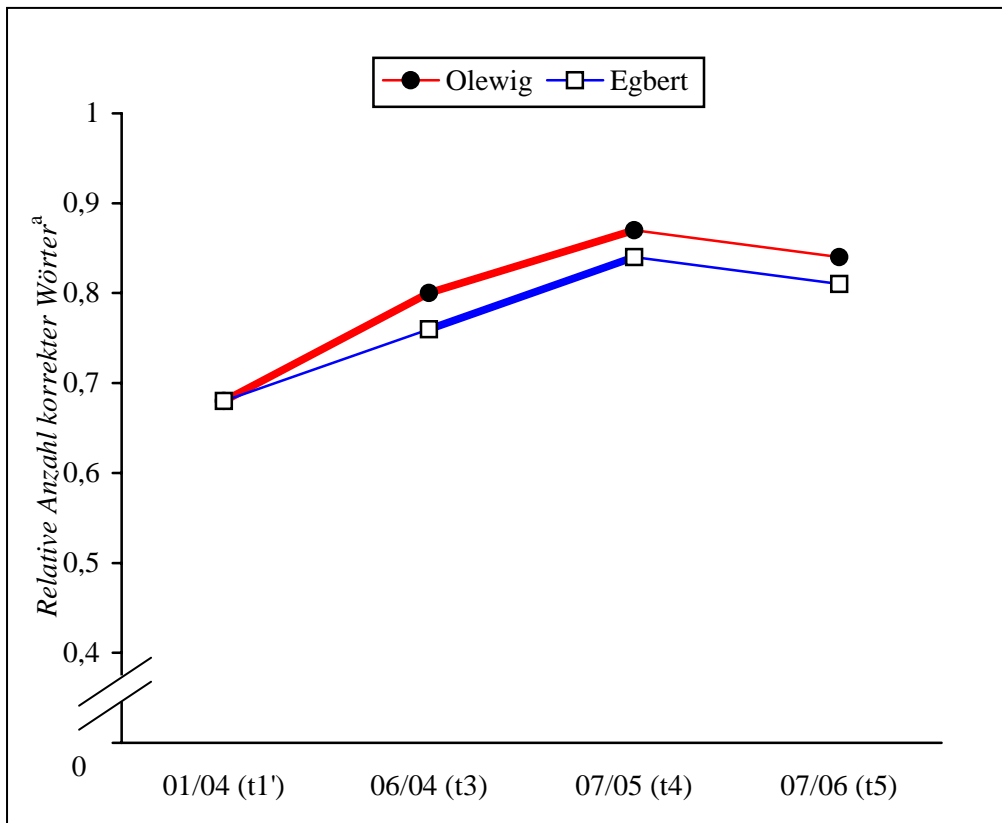


Abbildung 8.29 Relative Anzahl korrekt geschriebener Wörter im HSP über die vier Messzeitpunkte (t_1' – t_5) für Gruppe 2. Die Erhebung zu t_1' wurde verzögert durchgeführt, um diese zeitgleich mit der Erhebung des ersten Schuljahres stattfinden zu lassen.

^aDa zu den einzelnen Messzeitpunkten Testhefte mit unterschiedlicher Gesamtzahl von Wörtern verwendet wurden, sind aus Gründen der Vergleichbarkeit die an der Gesamtzahl der Wörter relativierten Häufigkeiten dargestellt.

8.2.4 Aspekte der schulischen Integration

Über die einzelnen Messzeitpunkte hinweg ist auf der Grundlage des *Fragebogens zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6)* für keine der beiden Schulen eine signifikante Veränderung auf der Dimension „Soziale Integration“ feststellbar. Auch im Vergleich der Schulen zueinander lässt sich zu keinem Messzeitpunkt ein bedeutsamer Unterschied feststellen. Indes zeigt sich im Vergleich der beiden Zeitpunkte t_3 und t_5 für die *emotionalen Integration* eine bedeutsame Abnahme in der Kontrollschule (Egbert). Das Ausmaß der emotionalen Integration zu t_4 unterscheidet sich nicht signifikant von den Werten für emotionale Integration zu t_3 oder t_5 . Der Vergleich der Schulen zueinander erbringt zu keinem Messzeitpunkt einen signifikanten Unterschied (vgl. Abb. 8.30).

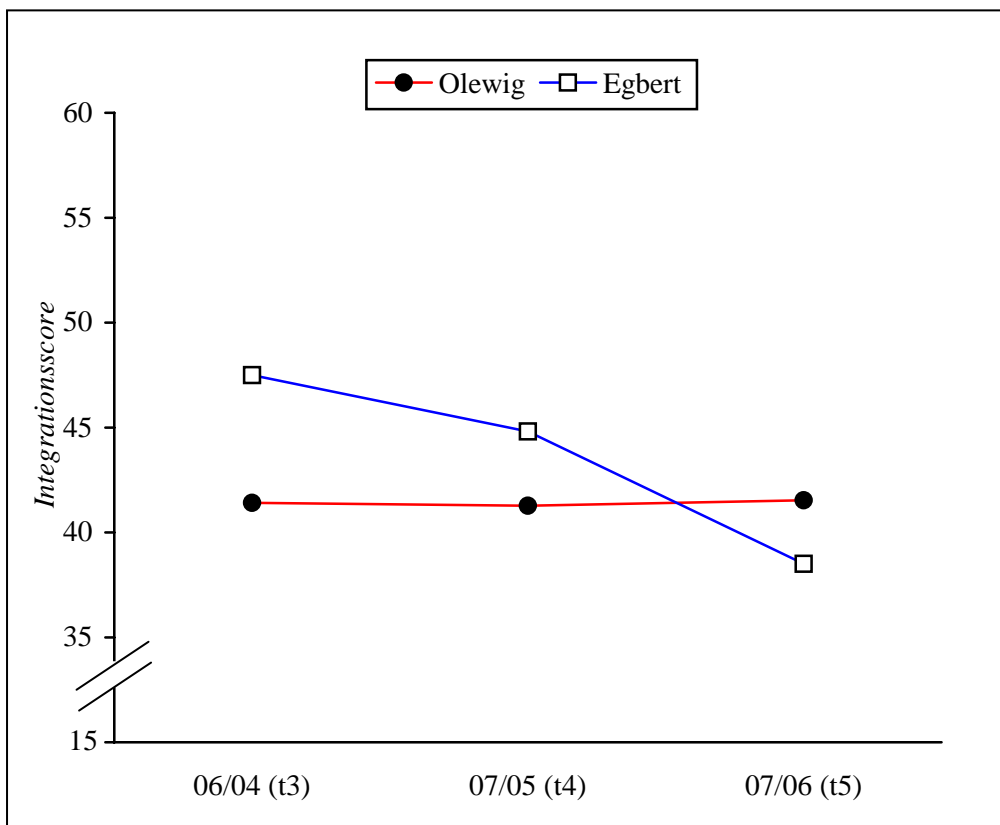


Abbildung 8.30 Scores für emotionale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 2 (theoretischer Range: 15 bis 60).

Auf Seiten der zu t_3 sozial nur schlecht integrierten Schülerinnen und Schüler zeigt sich ein bedeutsamer Anstieg der Integration von t_3 nach t_4 in der Experimentalschule (Olewig). In der Teilstichprobe der gut integrierten Schülerinnen und Schüler zeigt sich dagegen ein signifikanter Rückgang der sozialen Integration von t_3 nach t_4 in der Experimentalschule. Diesem folgt jedoch ein bedeutsamer Anstieg der Integrationswerte von t_4 nach t_5 . Zu t_3 liegen die Werte der Experimentalschule auf der Dimension „Soziale Integration“ bedeutsam über den Werten der Kontrollschule (Egbert). Für die Schülerinnen und Schüler der Kontrollschule lassen sich keine substantiellen Veränderungen über die Zeit nachweisen (vgl. Abb. 8.31).

Auf der Dimension „*Leistungsmotivationale Integration*“ hingegen zeigt sich von t_4 nach t_5 ein signifikanter Anstieg der Werte für die Schülerinnen und Schüler mit Schachunterricht (Olewig). Im Vergleich der beiden Schulen zueinander zeigen sich keine bedeutsamen Unterschiede (vgl. Abb. 8.32). Dieser Anstieg geht vor allem zurück auf die Teilstichprobe der Schülerinnen und Schüler, die anfänglich im Vergleich zu ihren Klassenkameraden geringe Werte auf der Dimension „*Leistungsmotivationale Integration*“ aufgewiesen hatten. Damit erzielt diese Teilstichprobe in Olewig zu t_5 auch signifikant höhere Werte als die in Egbert (vgl. Abb. 8.33).

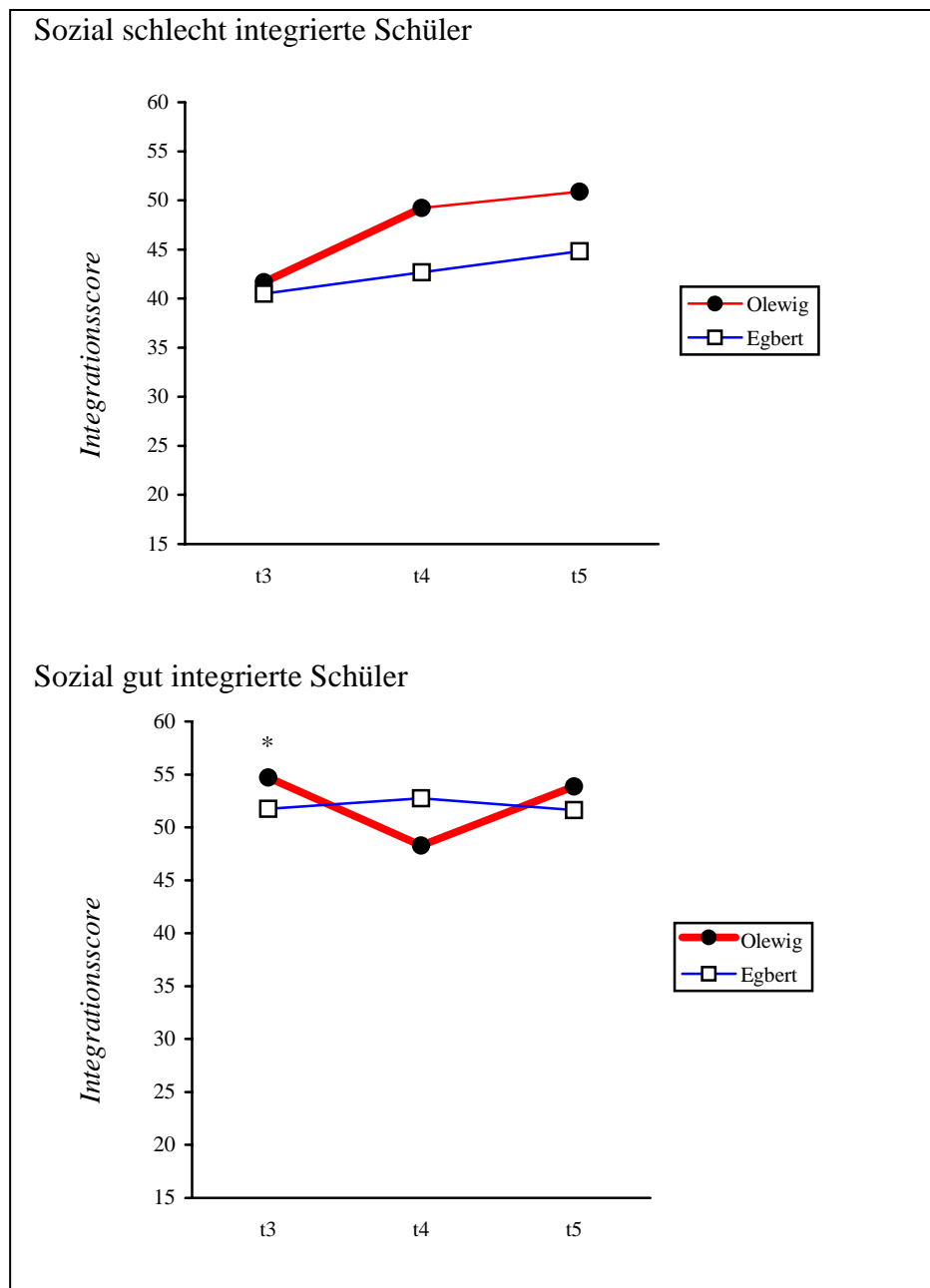


Abbildung 8.31 Scores für soziale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für wenig integrierte resp. gut integrierte Schüler aus Gruppe 2 (theoretischer Range: 15 bis 60).

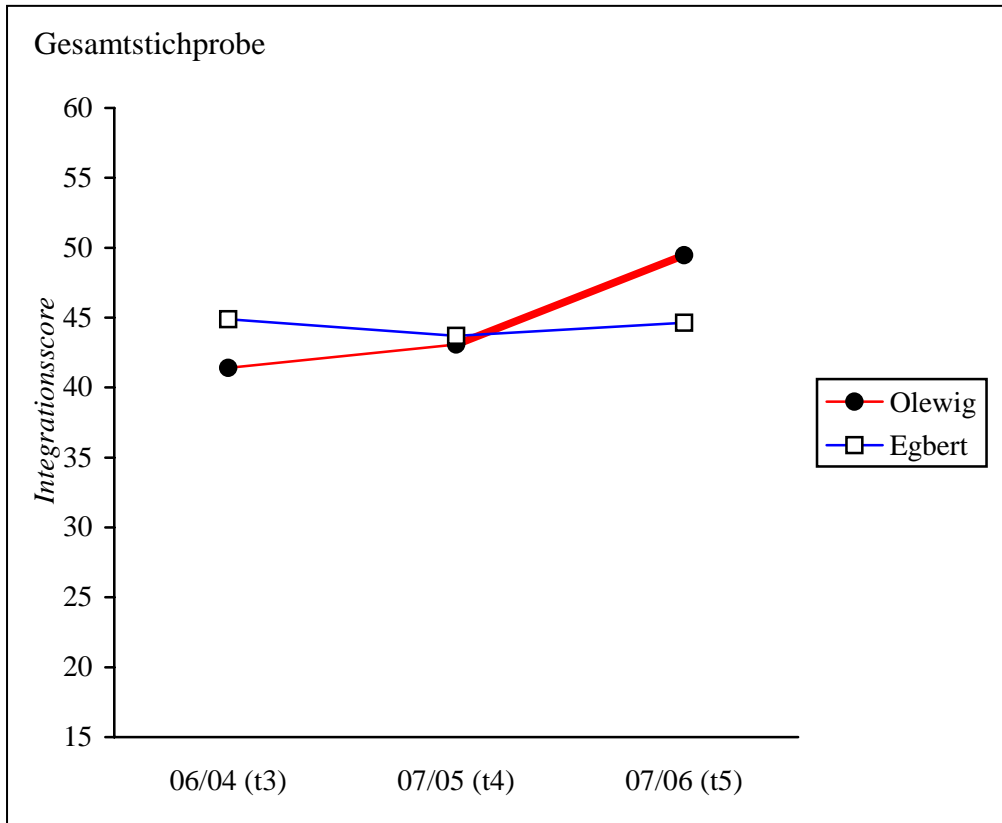


Abbildung 8.32 Scores für leistungsmotivationale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für Gruppe 2 (theoretischer Range: 15 bis 60).

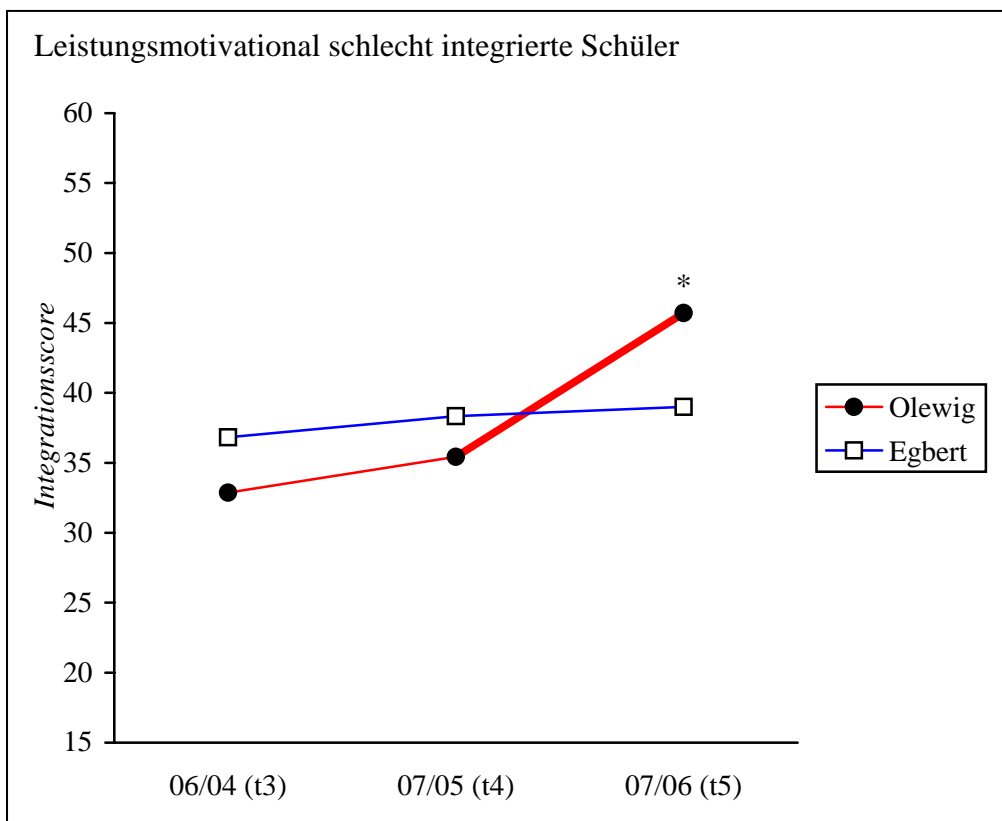


Abbildung 8.33 Scores für leistungsmotivationale Integration in die Schule über die drei Messzeitpunkte ($t_3 - t_5$) für schlecht integrierte Schüler aus Gruppe 2 (theoretischer Range: 15 bis 60).

8.3 Befunde für Gruppe 3

Die Schülerinnen und Schüler der Gruppe 3 besuchten zu Beginn der Evaluationsstudie das dritte Schuljahr. Nach dem Schuljahr 2004/2005 wechselten diese auf andere Schulen, so dass sie für die letzte Datenerhebung im Juli 2006 nicht mehr erreichbar waren. Die Schülerinnen und Schüler der Experimentalschule (Olewig) aus dieser Gruppe erhielten bis zu ihrem Schulwechsel zwei Jahre lang Schachunterricht.

8.3.1 Niveau der allgemeinen Intelligenz

Die Kinder beider Schulen verzeichnen bei der Bearbeitung der *Grundintelligenztestskala 20 (CFT 20)* gleichermaßen einen signifikanten Leistungsanstieg von t_3 nach t_4 , ohne dass sich zwischen den Schulen Unterschiede ergäben (vgl. Abb. 8.34).

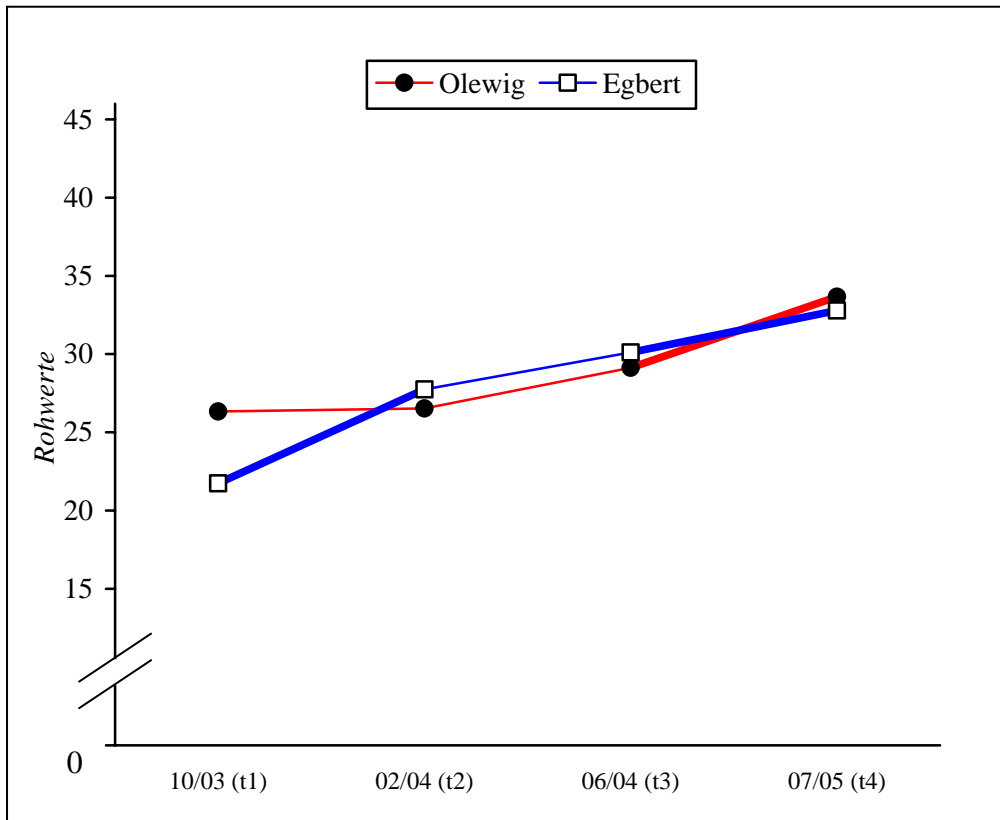


Abbildung 8.34 Gesamtscore der allgemeinen Intelligenz im CFT20 über die vier Messzeitpunkte für Gruppe 3 (theoretischer Range: 0 bis 46).

Für die Teilstichprobe der leistungsschwachen (nicht die der leistungsstarken) Kinder können signifikante Ergebnisse nachgewiesen werden. Diese liegen zum einen in einer bereits anfänglich (t_1) signifikant besseren Leistung der Experimentalschule und zum anderen in einem bedeutsamen Leistungsanstieg in der Kontrollschule von t_1 nach t_2 sowie einem ebenfalls bedeutsamen Leistungsanstieg in der Experimentalschule von t_3 nach t_4 (vgl. Abb. 8.35).

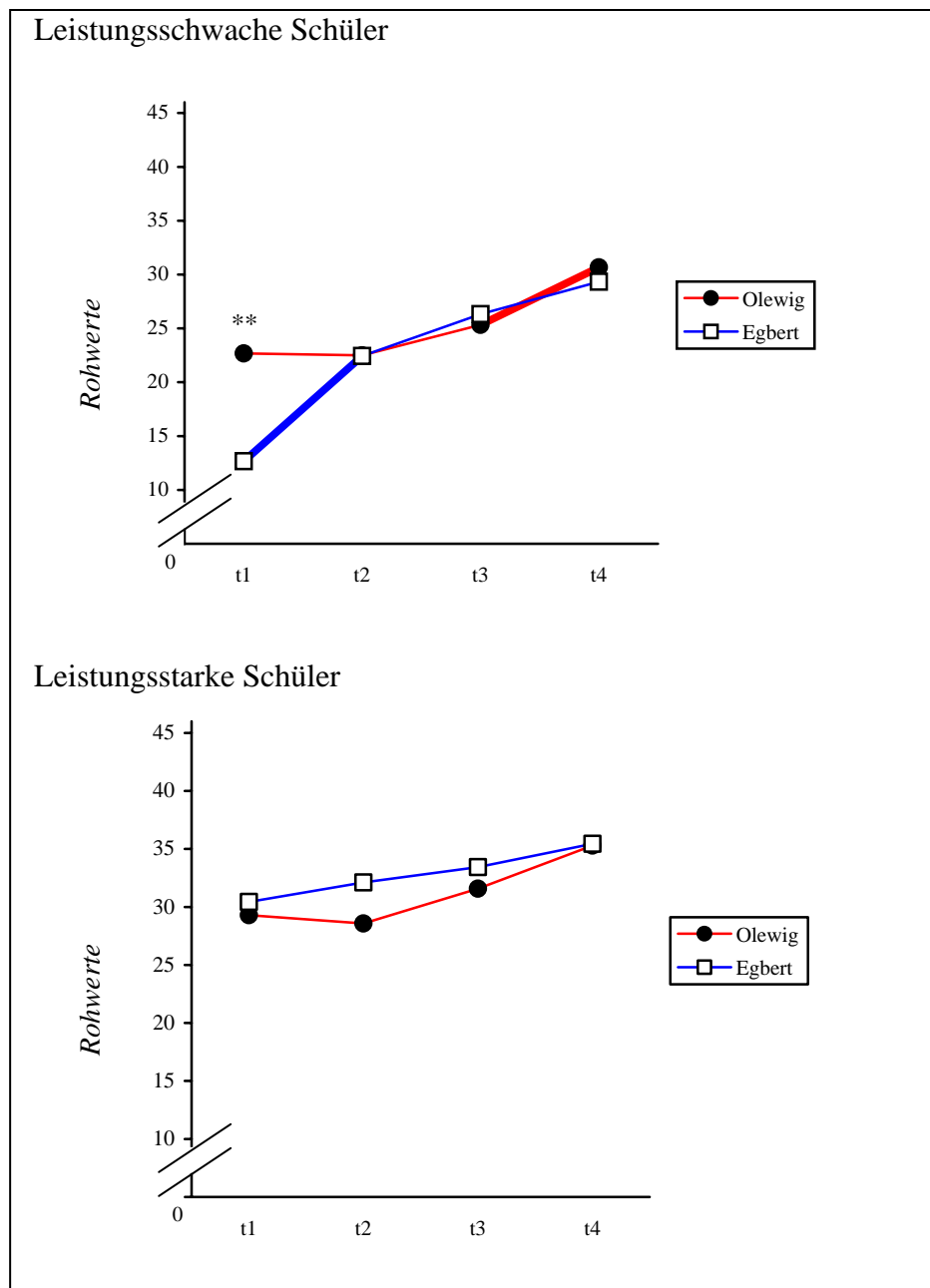


Abbildung 8.35 Gesamtscore der allgemeinen Intelligenz im CFT20 über die vier Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 3 (theoretischer Range: 0 bis 46).

Für den *Untertest* „Reihenfortsetzen“ der das Erkennen von Gesetzmäßigkeiten und deren Anwendung erfasst, zeigt sich ein signifikanter Zuwachs der Leistung in der Kontrollschule (Egbert) von t_1 nach t_2 . Der zu t_1 noch bestehende Leistungsunterschied zugunsten der Experimentalschule (Olewig) wird dadurch zu t_2 kompensiert (vgl. Abb. 8.36).

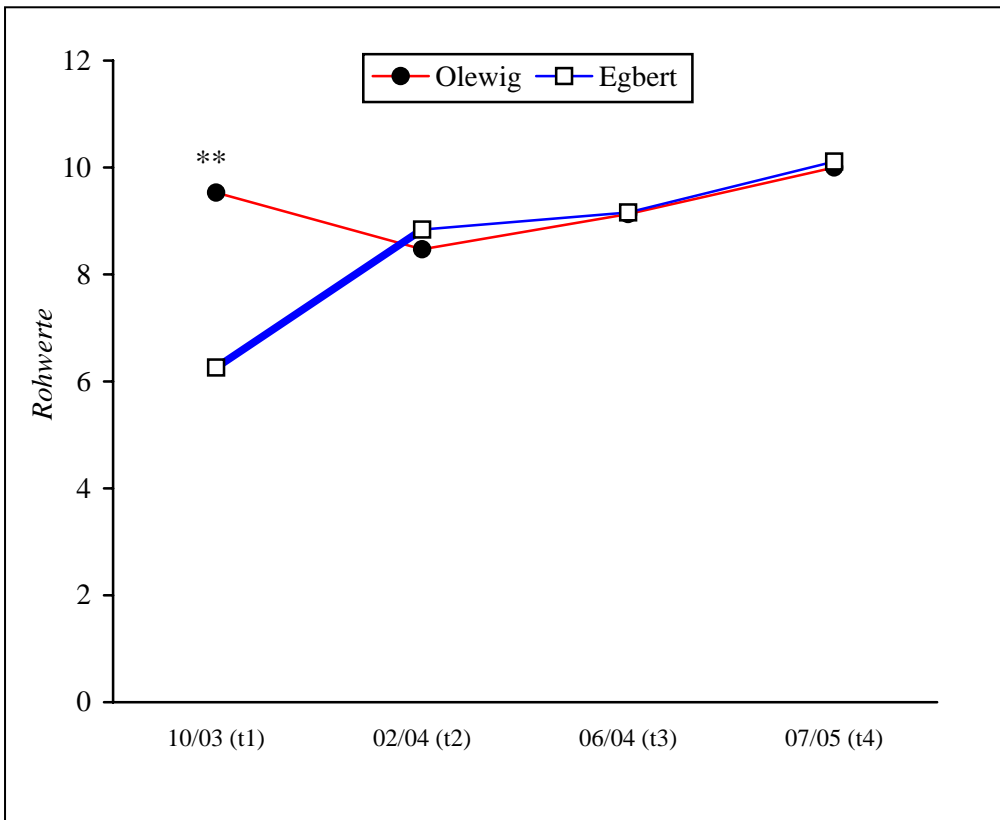


Abbildung 8.36 Absolute Rohwerte im Untertest „Reihenfortsetzen“ des CFT20 über die vier Messzeitpunkte für Gruppe 3 (theoretischer Range: 0 bis 12).

Die Analyse der Daten für die Teilstichprobe der leistungsschwachen Schülerinnen und Schüler liefert ein Ergebnismuster, das dem für die Gesamtstichprobe entspricht. Bei den leistungsstarken Schülerinnen und Schülern zeigt sich nur die Überlegenheit der Experimentalschule gegenüber der Kontrollschule zu t_1 ; die Veränderungen der Leistung über die Zeit erreichen keine statistische Signifikanz (vgl. Abb. 8.37).

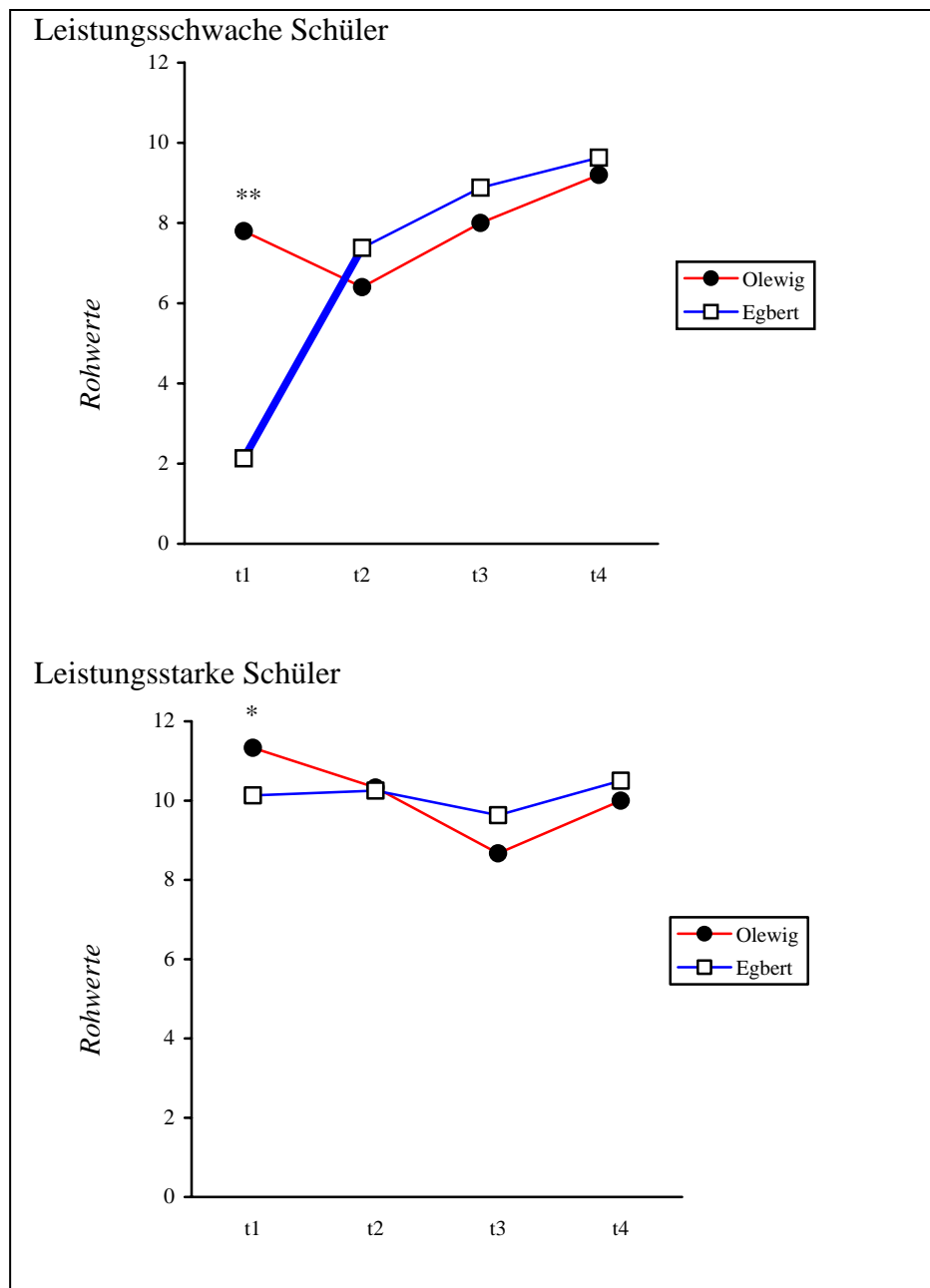


Abbildung 8.37 Absolute Rohwerte im Untertest „Reihenfortsetzen“ des CFT20 über die vier Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 3 (theoretischer Range: 0 bis 12).

Für den *Untertest* „Klassifikation“ (als Indikator des beziehungsstiftenden Denkens bei figuralem Material) zeigt sich als einzig signifikantes Ergebnis ein bedeutsamer Leistungsanstieg für die Schülerinnen und Schüler der Experimentalschule (Olewig) von t_3 nach t_4 (vgl. Abb. 8.38), der für die Kontrollschule nicht beobachtbar ist. Dieser Leistungsanstieg in der Gesamtstichprobe ist insbesondere auf die anfänglich leistungsschwachen Kinder zurückzuführen, während sich für die Gruppe der leistungsstarken Schülerinnen und Schüler keine signifikanten Veränderungen ergeben (vgl. Abb. 8.39).

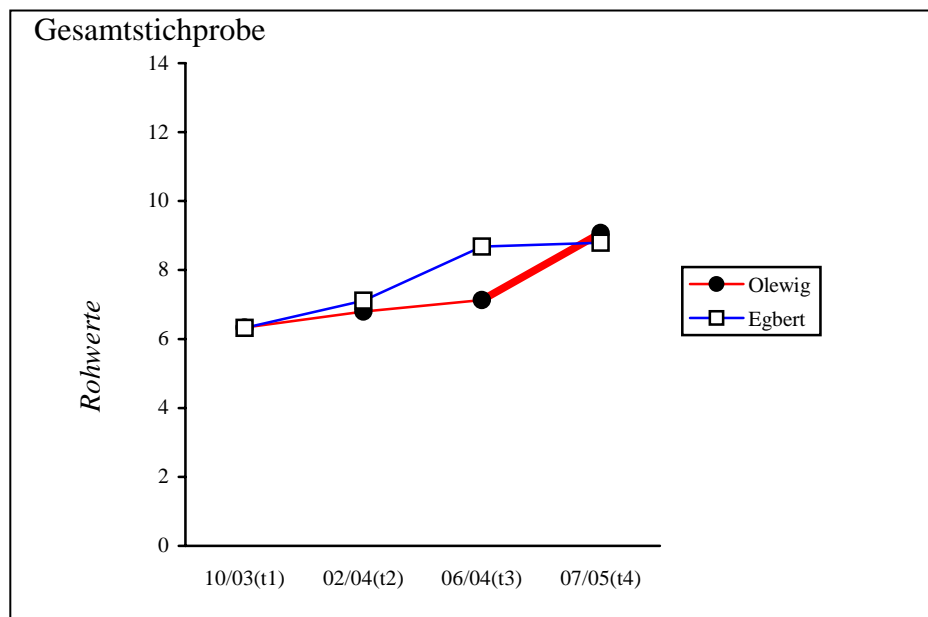


Abbildung 8.38 Absolute Rohwerte im Untertest „Klassifikation“ des CFT20 über die vier Messzeitpunkte für Gruppe 3 (theoretischer Range: 0 bis 14).

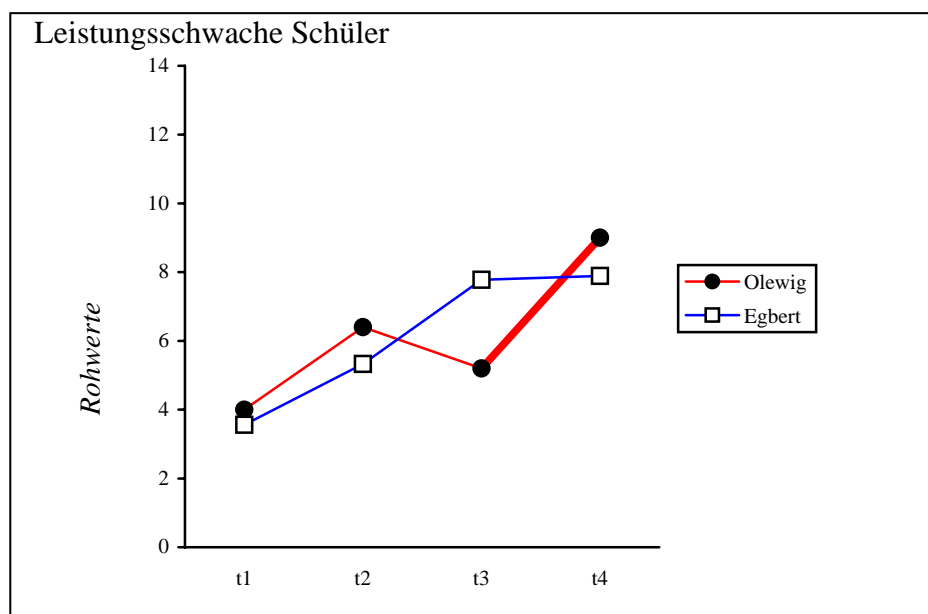


Abbildung 8.39 Absolute Rohwerte im Untertest „Klassifikation“ des CFT20 über die vier Messzeitpunkte für leistungsschwache Schüler aus Gruppe 3 (theoretischer Range: 0 bis 14).

8.3.2 Konzentrationsfähigkeit

Im *Differentiellen Leistungstest (DL-KG)* bestehen zu keinem der vier Messzeitpunkte Unterschiede in der Höhe der Konzentrationsfähigkeit zwischen den beiden Schulen, das gilt auch bei getrennter Betrachtung der leistungsstarken vs. leistungsschwachen Kinder beider Schulen. Gleichwohl ist ein signifikanter Anstieg der Konzentrationsfähigkeit von t_1 nach t_2 bei den Schülerinnen und Schülern mit Schachunterricht (Olewig) erkennbar, der sich auch für die Schülerinnen und Schüler ohne Schachunterricht (Egbert) zeigt. In der Kontrollschule zeigt sich darüber hinaus ein bedeutsamer Anstieg der Konzentrationsfähigkeit von t_2 nach t_3 (vgl. Abb. 8.40).

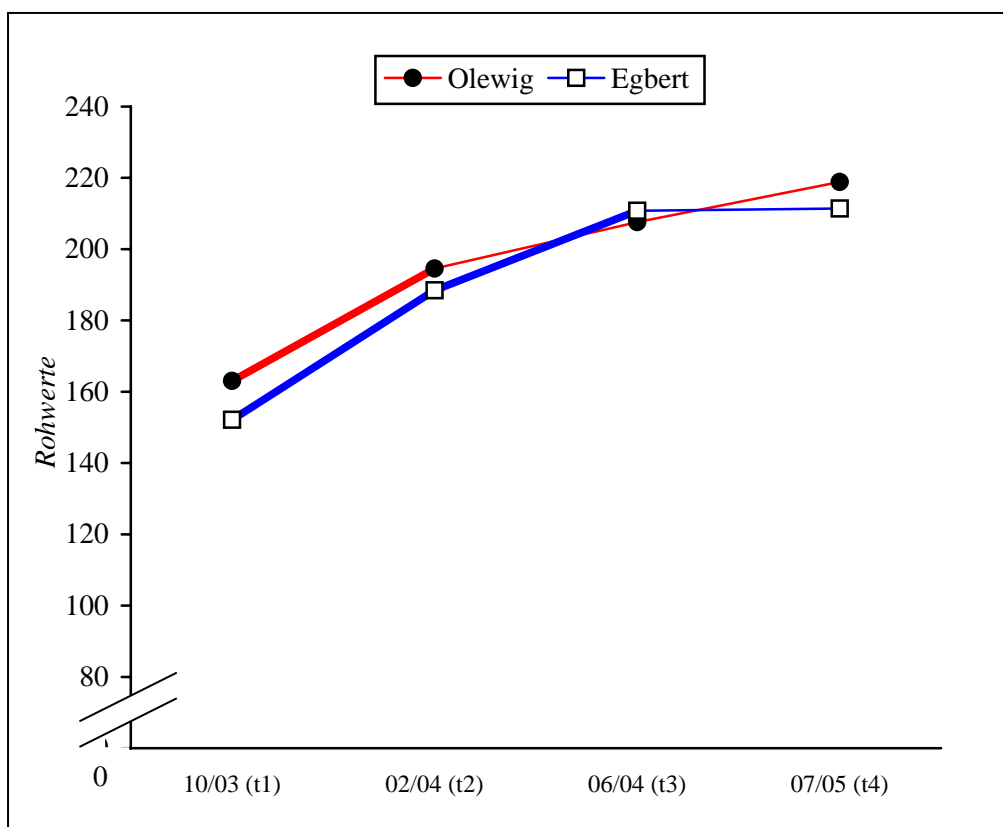


Abbildung 8.40 Absolute Rohwerte im DL-KG über die vier Messzeitpunkte (t_1 - t_4) für Gruppe 3 (theoretischer Range: 0 bis 240).

Bei gesonderter Betrachtung der leistungsschwachen Kinder zeigt sich in der Kontrollschule (Egbert) ein bedeutsamer Leistungsanstieg von t_1 nach t_2 und von t_2 nach t_3 . Für die entsprechenden Kinder der Experimentalschule (Olewig) zeigt sich ein solcher nicht. Die leistungsstarken Kinder beider Schulen zeigen einen bedeutsamen Leistungszuwachs von t_1 nach t_2 . (vgl. Abb. 8.41).

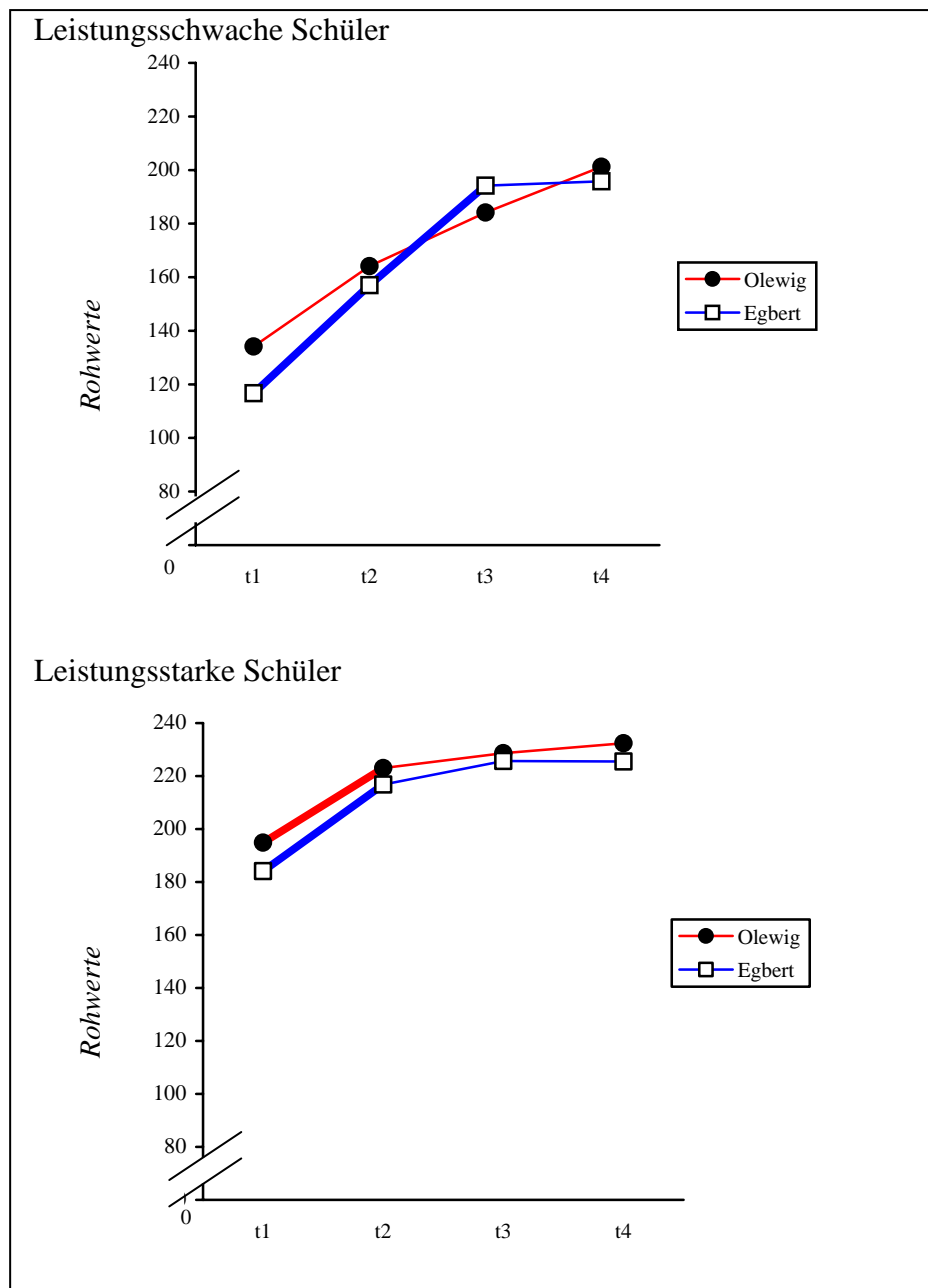


Abbildung 8.41 Absolute Rohwerte im DL-KG über die vier Messzeitpunkte ($t_1 - t_4$) für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 3 (theoretischer Range: 0 bis 240).

Die Kinder beider Schulen verzeichnen einen signifikanten Leistungsanstieg bei der Gesamtzahl abgeschriebener Silben in der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* von t_3 nach t_4 .

Zu t_3 sind beide Schulen noch nicht bedeutsam voneinander verschieden, so zeigt sich zu t_4 eine signifikant höhere Leistung der Schülerinnen und Schüler mit Schachunterricht (Olewig, vgl. Abb. 8.42).

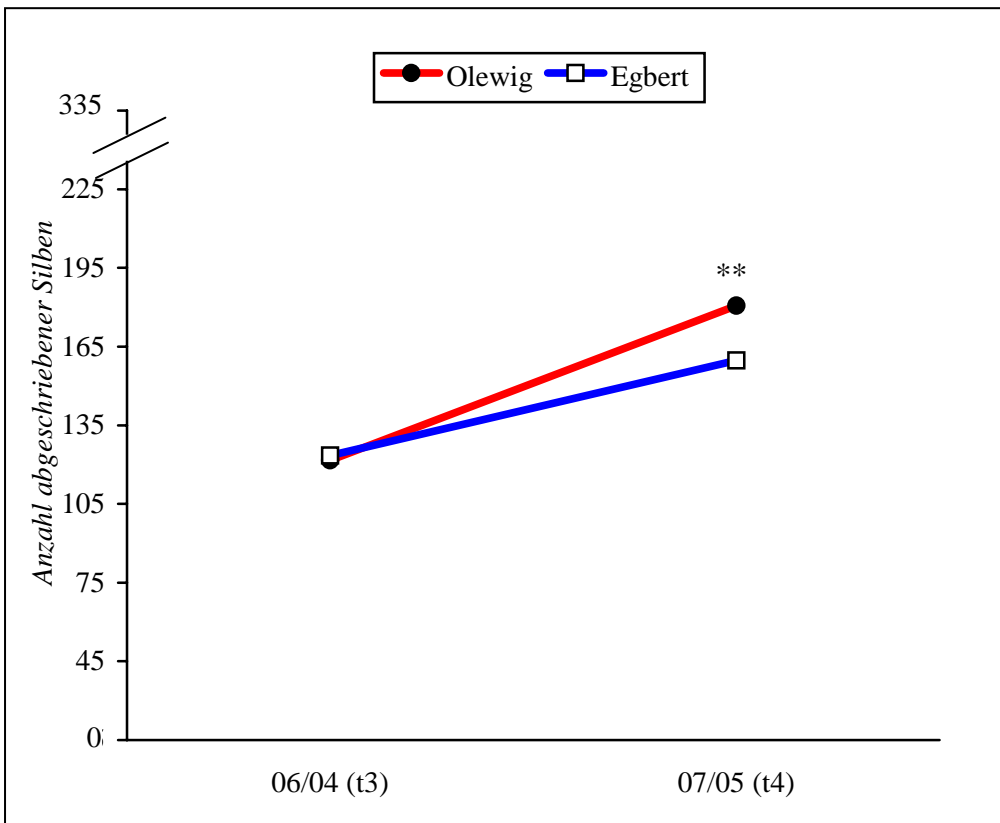


Abbildung 8.42 Anzahl abgeschriebener Silben im TPK über die beiden Messzeitpunkte (t_3 und t_4) für Gruppe 3 (theoretischer Range: 0 bis 335).

Die Anzahl der beim Abschreiben gemachten Fehler verändert sich in keiner der beiden Gruppen. Während sich zu t_1 sich kein Unterschied zwischen den beiden Schulen findet, zeigt sich eine signifikant geringere Fehlerzahl in der Kontrollschule (Egbert) zu t_2 (vgl. Abb. 8.43).

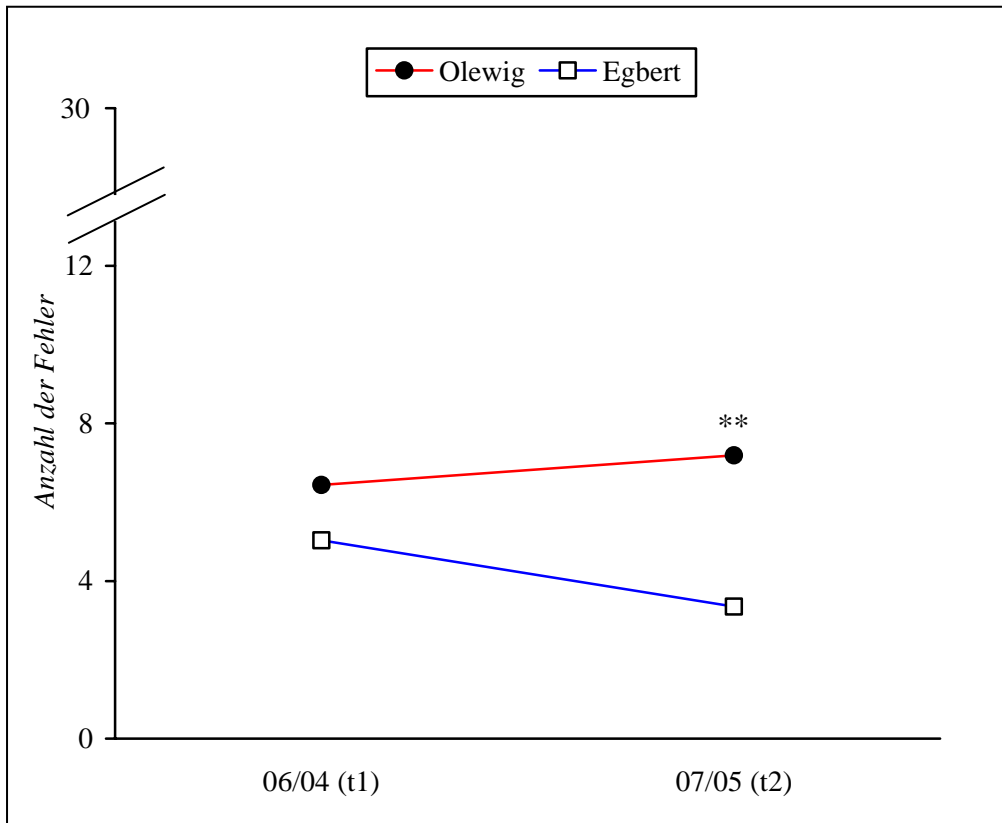


Abbildung 8.43 Anzahl der Fehler im Abschreibtext des TPK über die beiden Messzeitpunkte (t_1 und t_2) für Gruppe 3.

Lediglich die Kinder der Kontrollschule (Egbert) weisen einen statistisch bedeutsamen Anstieg der rezeptiven Aufmerksamkeitsleistung von t_3 nach t_4 auf. Der Vergleich der beiden Schulen zueinander ergibt zu keinem der beiden Messzeitpunkte bedeutsame Unterschiede (vgl. Abb. 8.44).

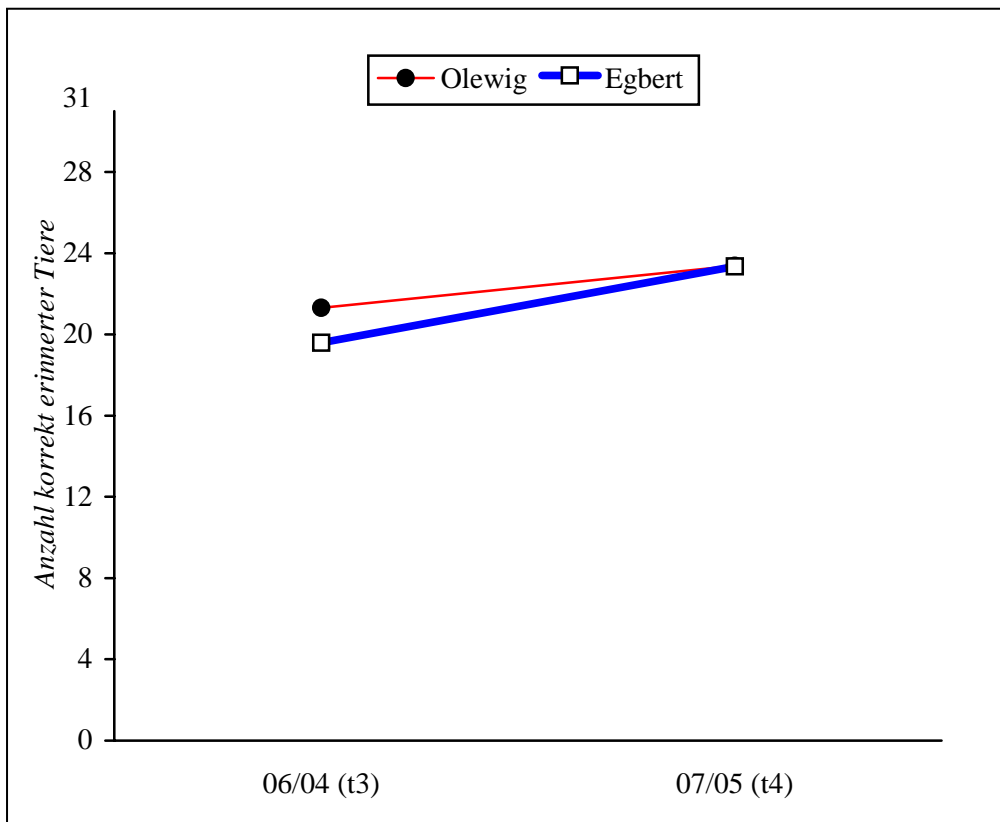


Abbildung 8.44 Anzahl korrekt erinnerter Tiere im TPK über die beiden Messzeitpunkte (t_3 und t_4) für Gruppe 3 (theoretischer Range: 0 bis 31).

8.3.3 Prüfung von Rechen- und Rechtschreibleistungen

Die Analyse der in der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* bearbeiteten Rechenaufgaben ergibt für keine der beiden Schulen eine bedeutsame Leistungsveränderung zwischen den beiden Messzeitpunkten; das gilt sowohl für die Gesamtzahl der bearbeiteten Rechenaufgaben als auch für die Fehlerzahl. Im Vergleich der beiden Schulen zueinander bearbeitet die Experimentalschule (Olewig) zu t_4 signifikant mehr Aufgaben als die Kontrollschule (Egbert, vgl. Abb. 8.45).

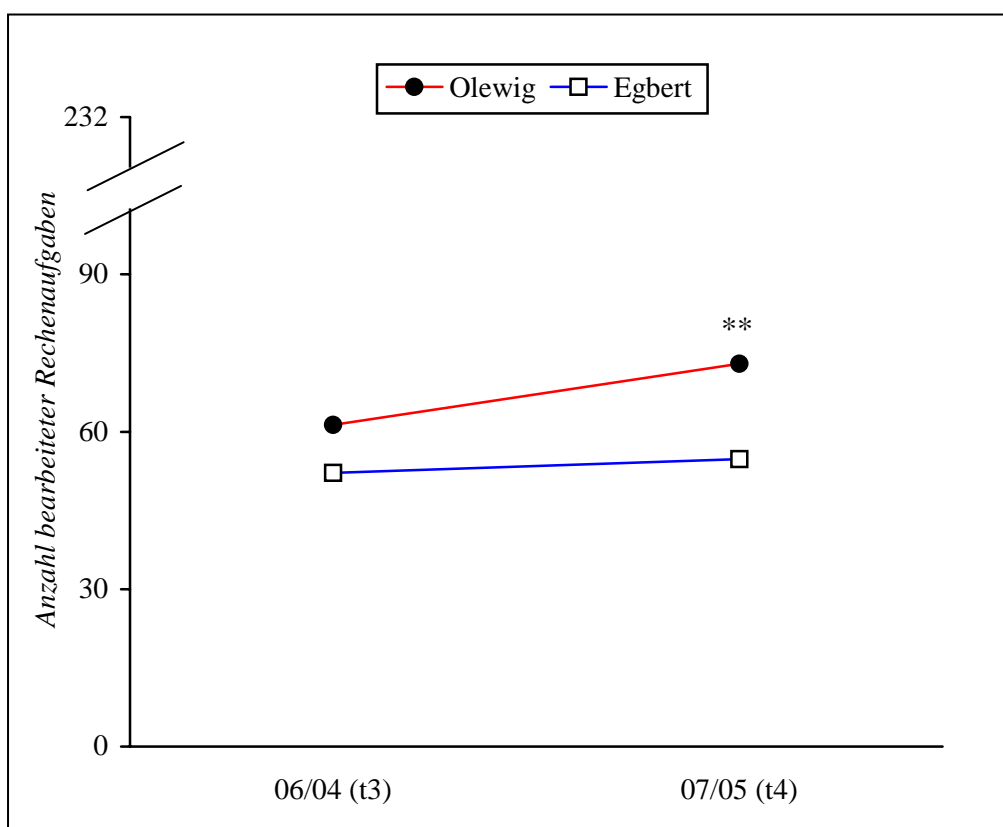


Abbildung 8.45 Anzahl bearbeiteter Rechenaufgaben im TPK über die beiden Messzeitpunkte (t_3 und t_4) für Gruppe 3 (theoretischer Range: 0 bis 232).

In der *Hamburger Schreibprobe (HSP)* zeigen die Kinder beider Schulen – und zwar insbesondere die leistungsschwachen – von t_1' nach t_3 eine bedeutsame Verbesserung der Rechtschreibleistung. Von t_3 nach t_4 können lediglich jene mit Schachunterricht (Olewig) einen weiteren bedeutsamen Leistungsanstieg verzeichnen (vgl. Abb. 8.46), der insbesondere auf einen signifikanten Anstieg bei den leistungsstarken Kindern zurückgeht. Der Vergleich der Schulen zueinander ergibt keine signifikanten Unterschiede zu einem der Messzeitpunkte, wenn man die Gesamtklassen betrachtet; allerdings erzielten unter den leistungsstarken Schülerinnen und Schülern jene aus der Kontrollschule (Egbert) zu den Messzeitpunkten t_1' und t_3 signifikant höhere Ergebnisse als die leistungsstarken Kinder der Experimentalschule (Olewig).

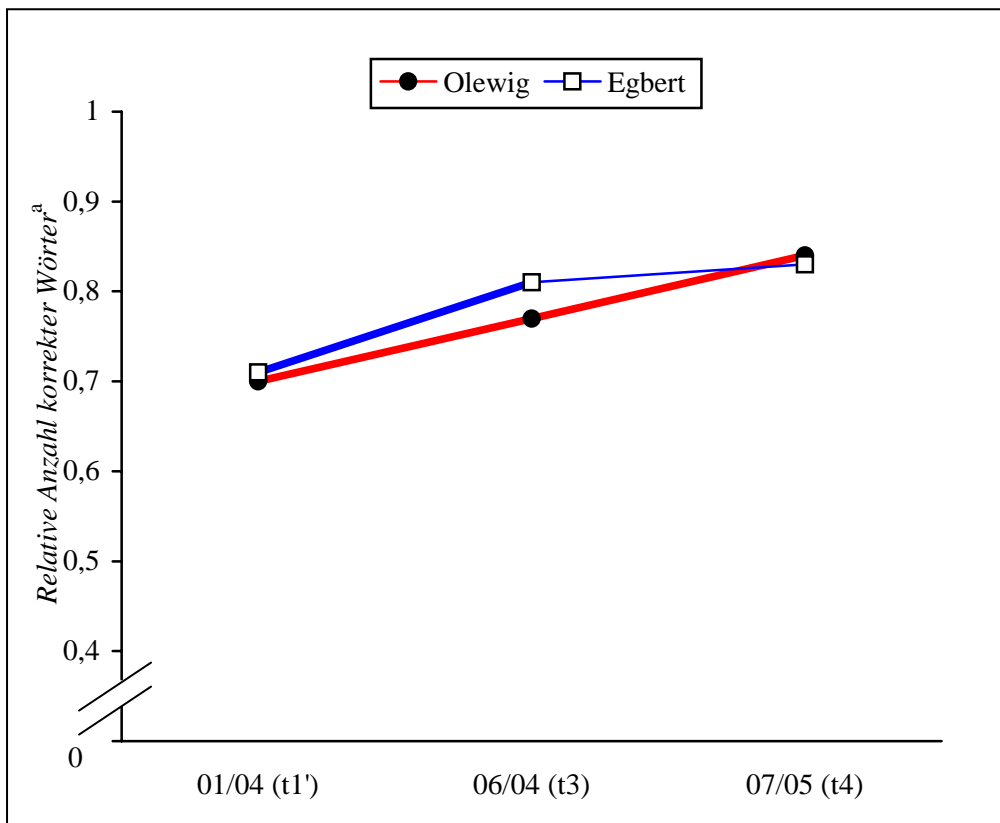


Abbildung 8.46 Relative Anzahl korrekter Wörter im HSP über die drei Messzeitpunkte (t_1' – t_4) für Gruppe 3. Die Erhebung zu t_1' wurde verzögert durchgeführt, um diese zeitgleich mit der Erhebung des ersten Schuljahres stattfinden zu lassen.

^aDa zu den einzelnen Messzeitpunkten Testhefte mit unterschiedlicher Gesamtzahl von Wörtern verwendet wurden, sind aus Gründen der Vergleichbarkeit die an der Gesamtzahl der Wörter relativierten Häufigkeiten dargestellt.

8.3.4 Aspekte der schulischen Integration

Im *Fragebogen zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6)* verändert sich der Wert für „*Soziale Integration*“ in keiner der Schulen von t_3 nach t_4 bedeutsam. Auch der Vergleich der Schulen zueinander zum zweiten Messzeitpunkt ergibt keinen signifikanten Unterschied. Das einzige statistisch bedeutsame Ergebnis an dieser Stelle stellt der von t_3 nach t_4 in der Kontrollschule (Egbert) zu verzeichnende Anstieg der sozialen Integration auf Seiten der sozial schlecht integrierten Schülerinnen und Schüler dar.

Der Wert für „*Emotionale Integration*“ verändert sich von t_3 nach t_4 in keiner der beiden Schulen bedeutsam. Auch im Vergleich der beiden Schulen zueinander sind zu t_4 keine signifikanten Unterschiede festzustellen. Für die Teilstichprobe der Schülerinnen und Schüler mit relativ niedrigen Integrationsscores zu t_3 finden sich keine signifikanten Ergebnisse. Auf Seiten der Schülerinnen und Schüler mit (zu t_3) relativ hohen Werten auf der Dimension „*Emotionale Integration*“ findet sich im Vergleich der beiden Schulen zueinander ein bedeutsamer Unterschied zugunsten der Kontrollschule (Egbert) zum dritten Messzeitpunkt.

Auf der Dimension „*Leistungsmotivationale Integration*“ zeigt sich für keine der beiden Gruppen eine signifikante Veränderung zwischen den beiden Messzeitpunkten. Der Vergleich zwischen den beiden Schulen zeigt einen bedeutsamen Unterschied, wobei die Experimentalschule (Olewig) höhere Werte erzielt (vgl. Abb. 8.47).

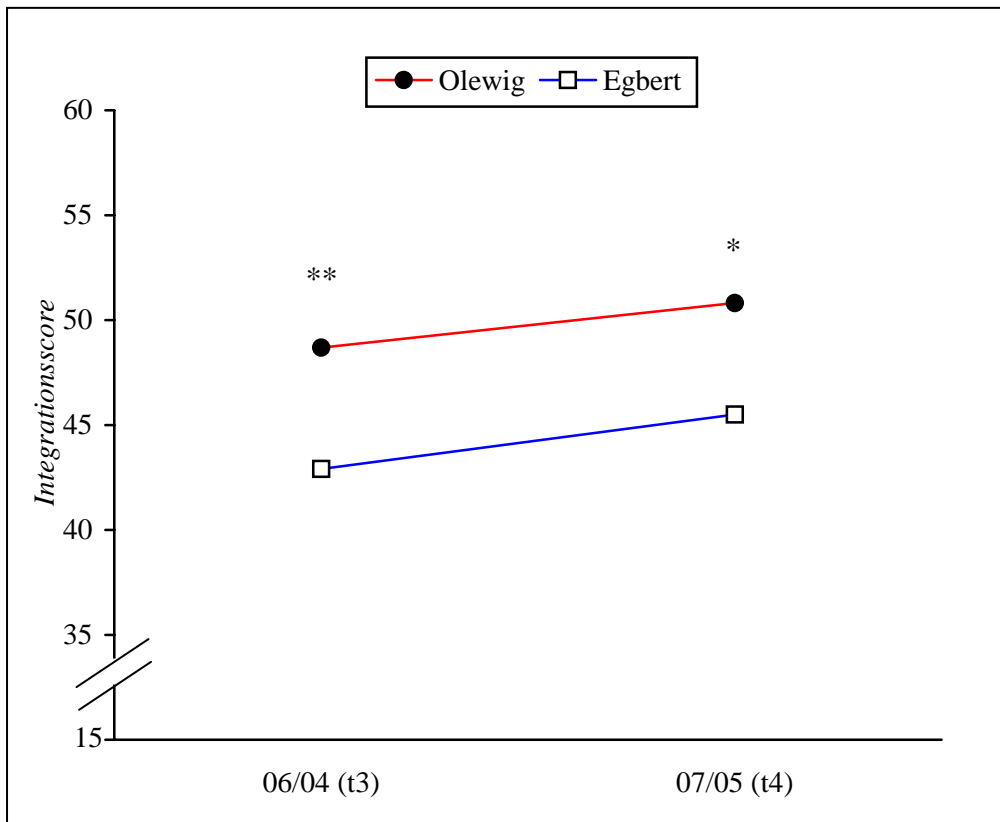


Abbildung 8.47 Scores für leistungsmotivationale Integration in die Schule über die beiden Messzeitpunkte ($t_3 - t_4$) für Gruppe 3 (theoretischer Range: 15 bis 60).

Die anfänglich weniger gut integrierten Schülerinnen und Schüler können in beiden Schulen einen ähnlich verlaufenden Anstieg der Integrationswerte von t_3 nach t_4 verzeichnen. Zu beiden Messzeitpunkten liegen die Werte der Experimentalschule (Olewig) signifikant über den Werten der Kontrollschule (Egbert). In der Teilstichprobe der hoch integrierten Schülerinnen und Schüler ist der Unterschied zwischen den beiden Gruppen nur zu t_3 signifikant, wobei sich keine bedeutsamen Veränderungen auf dieser Dimension zeigen (vgl. Abb. 8.48).

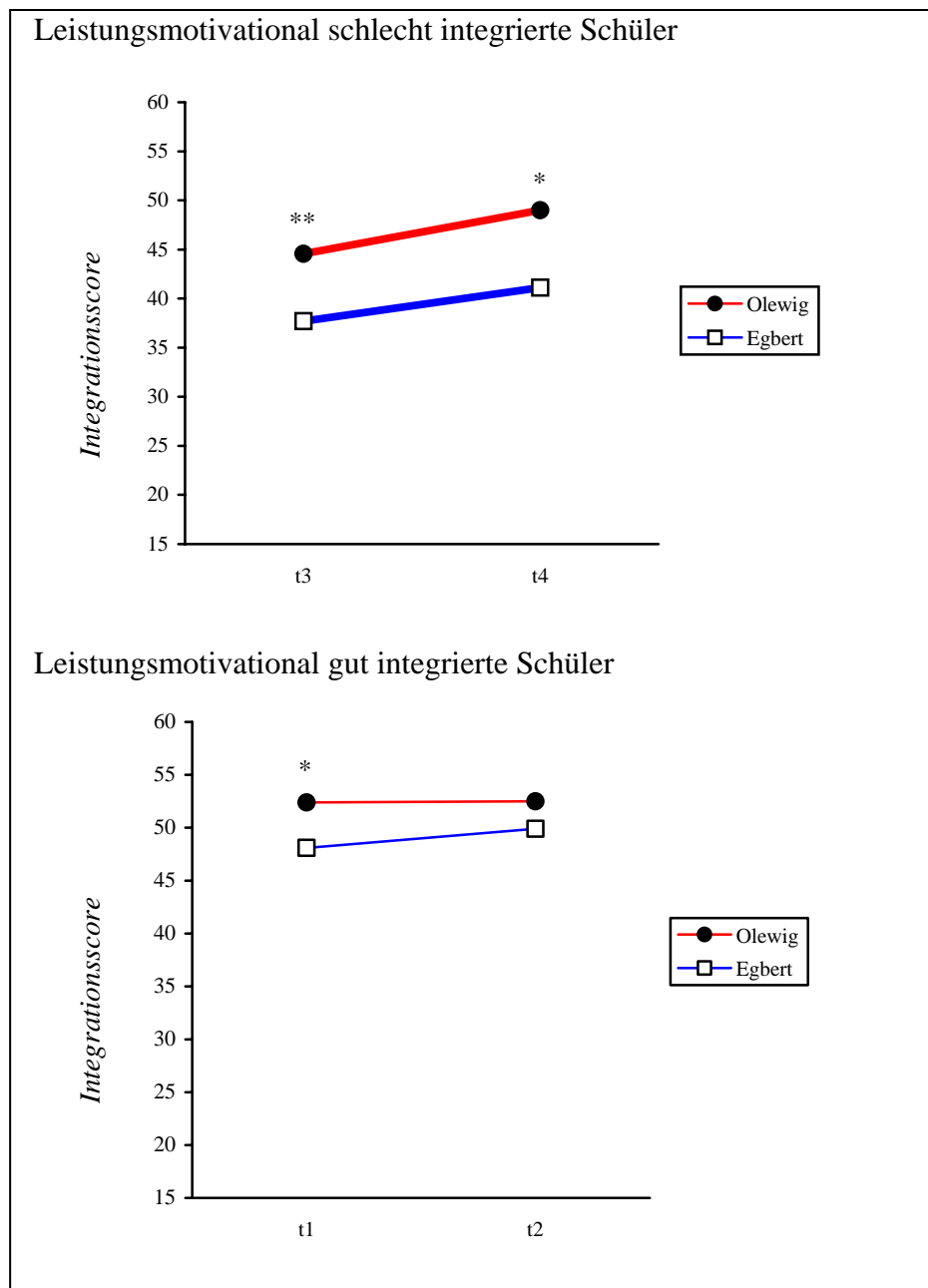


Abbildung 8.48 Scores für leistungsmotivationale Integration in die Schule über die beiden Messzeitpunkte ($t_3 - t_4$) für schlecht integrierte resp. gut integrierte Schüler aus Gruppe 3 (theoretischer Range: 15 bis 60).

8.4 Befunde für Gruppe 4

Die Schülerinnen und Schüler der Gruppe 4 besuchten zu Beginn der Evaluationsstudie das vierte Schuljahr. Da diese Kinder wegen des Wechsels auf weiterführende Schulen schon im Schuljahr 2004/2005 nicht mehr erreichbar waren, liegen für Gruppe 4 nur die Daten zu den ersten Messzeitpunkten innerhalb des Schuljahres 2003/2004 vor. Die Schülerinnen und Schüler der Experimentalschule (Olewig) aus dieser Gruppe erhielten bis zu ihrem Schulwechsel lediglich ein Jahr lang Schachunterricht.

8.4.1 Niveau der allgemeinen Intelligenz

Die Kinder beider Schulen unterscheiden sich in ihren Leistungen in der *Grundintelligenztestskala 20 (CFT 20)* zu keinem Messzeitpunkt bedeutsam voneinander; ein bedeutsamer Leistungsanstieg ist von t_2 nach t_3 zu verzeichnen, und zwar in beiden Schulen (vgl. Abb. 8.49).

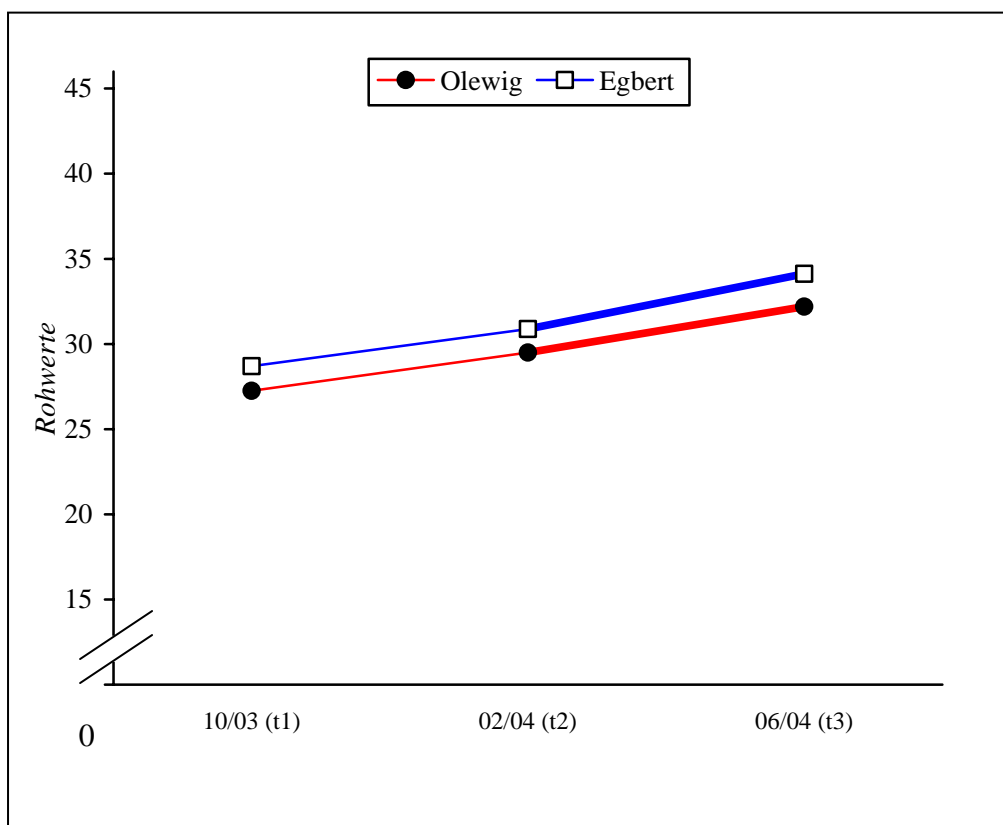


Abbildung 8.49 Gesamtscore der allgemeinen Intelligenz im CFT20 über die drei Messzeitpunkte für Gruppe 4 (theoretischer Range: 0 bis 46).

Für die leistungsschwachen Schülerinnen und Schüler zeigt sich ein signifikanter Zuwachs der Leistung von t_1 nach t_2 in der Experimentalschule (Olewig); ein solcher zeigt sich allerdings auch in der Kontrollschule (Egbert), hier zwischen t_2 und t_3 . In der Gruppe der leistungsstarken Schülerinnen und Schüler ist lediglich in der Experimentalschule ein bedeutsamer Anstieg der Leistung zu beobachten (von t_2 nach t_3 , vgl. Abb. 8.50).

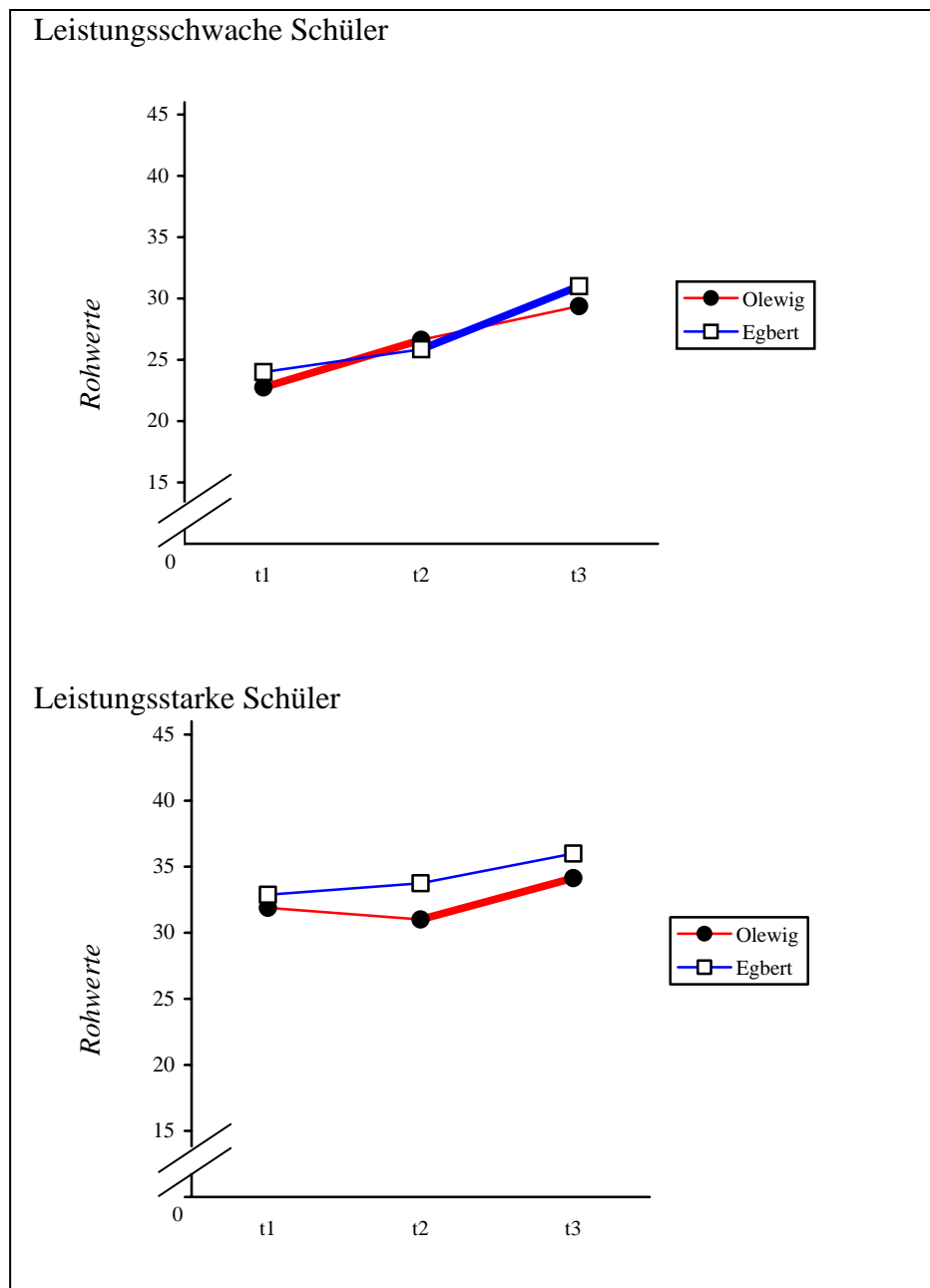


Abbildung 8.50 Gesamtscore der allgemeinen Intelligenz im CFT20 über die drei Messzeitpunkte für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 4 (theoretischer Range: 0 bis 46).

Zu t_1 und t_3 sind die Leistungen im *Untertest* „Reihenfortsetzen“ der Kontrollschule (Egbert) signifikant höher als die der Experimentalschule (Olewig). Eine bedeutsame Leistungsverbesserung zeigt sich nur in der Experimentalschule (Olewig), und zwar nur von t_1 nach t_2 , die im Wesentlichen auf entsprechende Verbesserungen bei den leistungsschwächeren Kindern zurückgeht und die den anfänglichen Unterschied zwischen beiden Schulen kompensiert (vgl. Abb. 8.51 und 8.52).

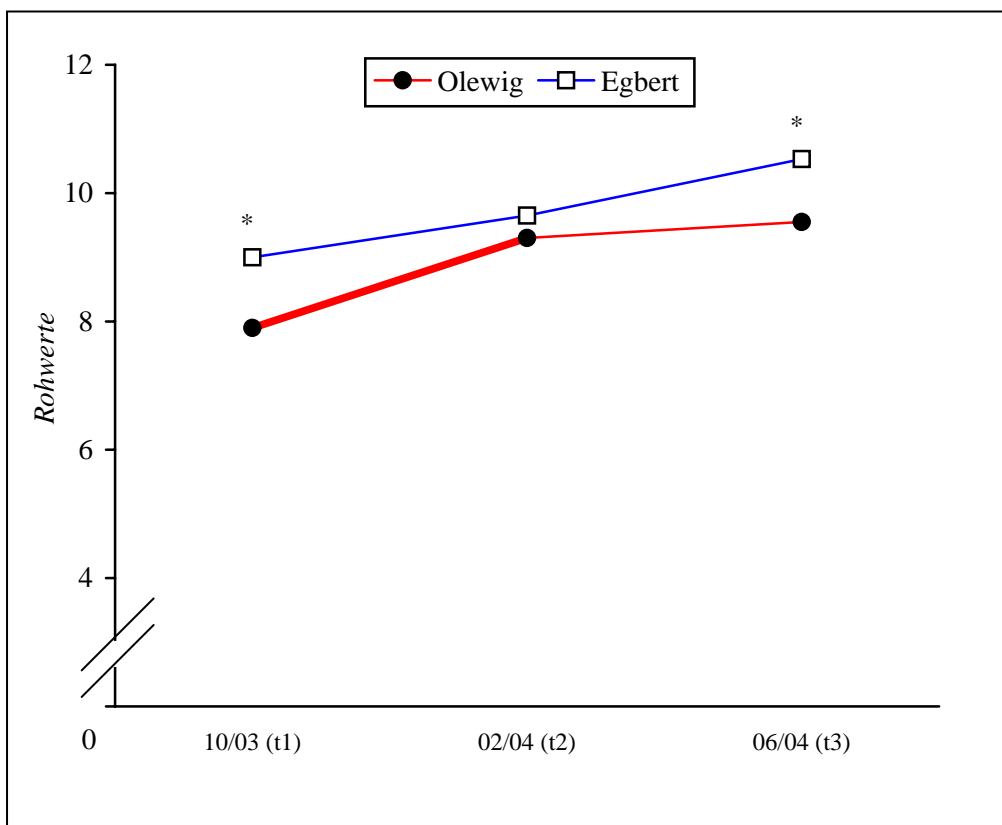


Abbildung 8.51 Absolute Rohwerte im Untertest „Reihenfortsetzen“ des CFT20 über die drei Messzeitpunkte für Gruppe 4 (theoretischer Range: 0 bis 12).

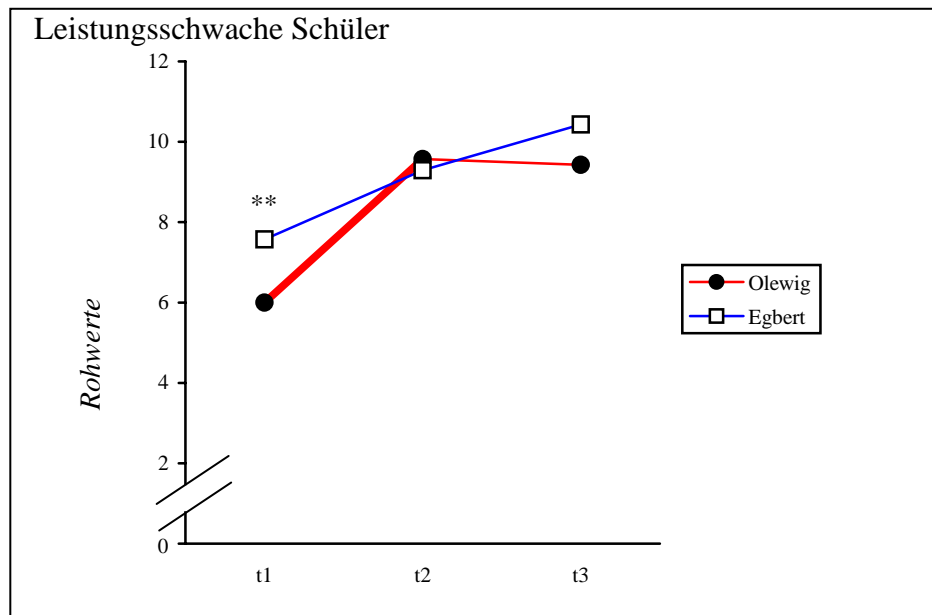


Abbildung 8.52 Absolute Rohwerte im Untertest „Reihenfortsetzen“ des CFT20 über die drei Messzeitpunkte für leistungsschwache resp. Schüler aus Gruppe 4 (theoretischer Range: 0 bis 12).

Auch bezogen auf *Untertest „Klassifikation“* zeigt sich für die Teilstichprobe der leistungsschwachen Schülerinnen und Schüler ein bedeutsamer Anstieg der Leistungen von t_1 nach t_2 , und zwar hier in beiden Schulen (vgl. Abb. 8.53).

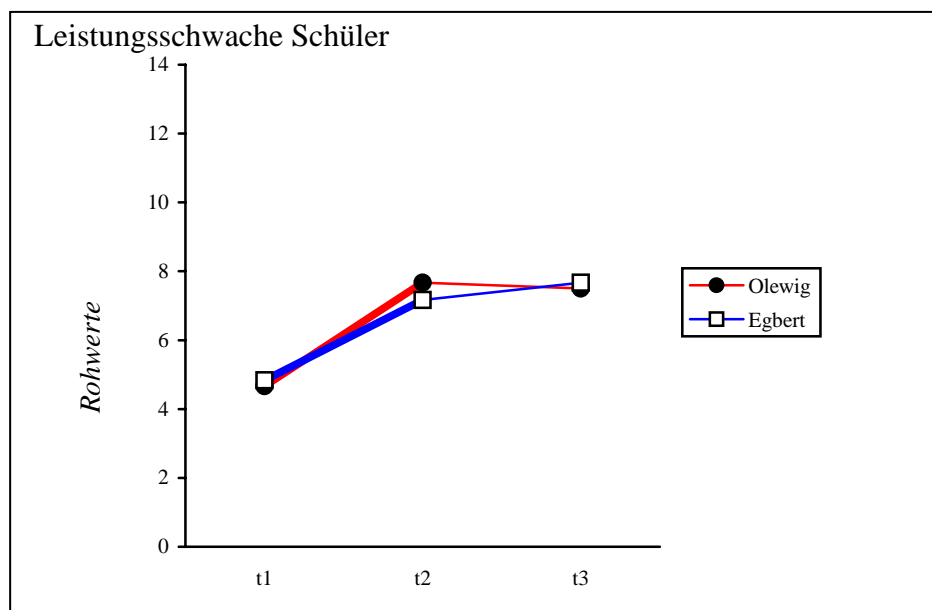


Abbildung 8.53 Absolute Rohwerte im Untertest „Klassifikation“ des CFT20 über die drei Messzeitpunkte für leistungsschwache Schüler aus Gruppe 4 (theoretischer Range: 0 bis 14).

Für den *Untertest „Topologisches Schlussfolgern“* (in dem die räumliche Lage verschiedener Elemente in einer Vorlage zueinander erkannt und auf die Lösungsalternativen übertragen werden muss) schließlich zeigt sich ein bedeutsamer Anstieg der Leistungen für die Experimentalschule von t_2 nach t_3 in der Gesamtstichprobe (vgl. Abb. 8.54).

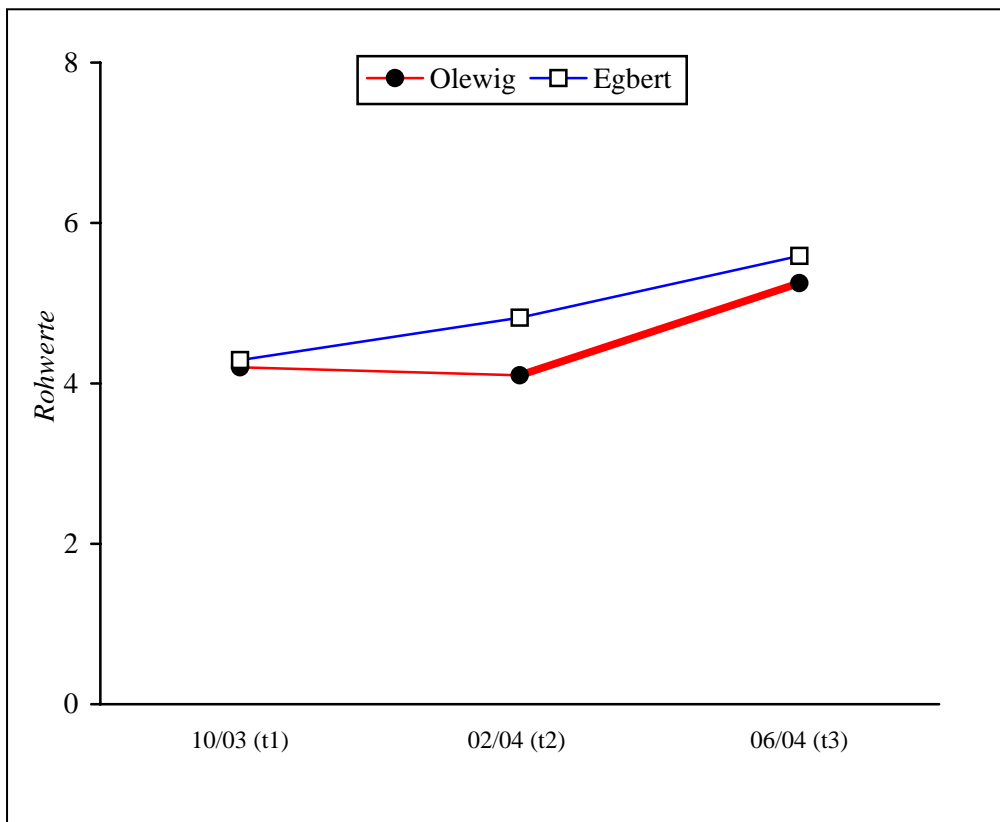


Abbildung 8.54 Absolute Rohwerte im Untertest „Topologisches Schlussfolgern“ des CFT20 über die drei Messzeitpunkte für Gruppe 4 (theoretischer Range: 0 bis 8).

8.4.2 Konzentrationsfähigkeit

Die Kinder beider Schulen verzeichnen im *Differentiellen Leistungstest (DL-KG)* einen signifikanten Leistungszuwachs von t_1 nach t_2 . Wie im 3. Schuljahr kann nur die Kontrollschule noch einen leichten Zuwachs von t_2 nach t_3 erzielen und den (anfänglich höheren) Wert der Experimentalschule erreichen (vgl. Abb. 8.55).

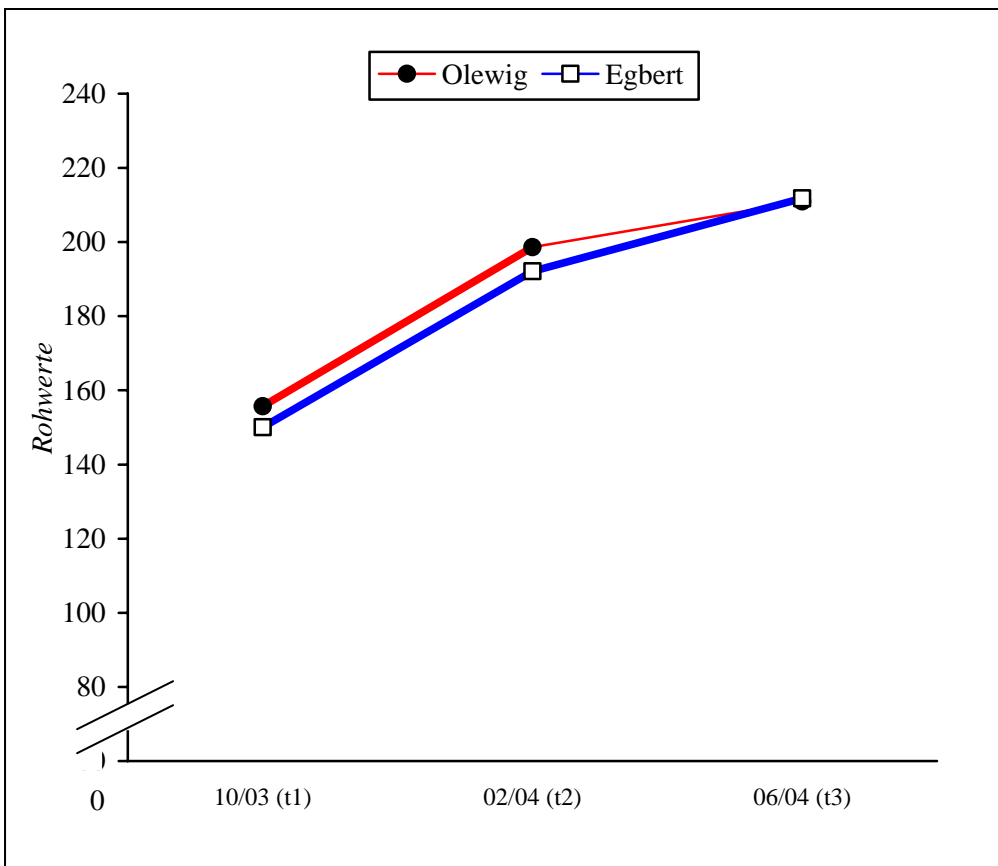


Abbildung 8.55 Absolute Rohwerte im DL-KG über die drei Messzeitpunkte (t_1 - t_3) für Gruppe 4 (theoretischer Range: 0 bis 240).

Auf Seiten der leistungsschwachen, aber auch bei den leistungsstarken Schülerinnen und Schülern zeigt sich ein signifikanter Anstieg der Leistung von t_1 nach t_2 in beiden Schulen. Darüber hinaus zeigt sich für die leistungsstarken Schülerinnen und Schüler der Kontrollschule (Egbert) ein Anstieg auch noch von t_2 nach t_3 . Signifikante Unterschiede zwischen den Schulen bestehen zu keinem der Messzeitpunkte (vgl. Abb. 8.56).

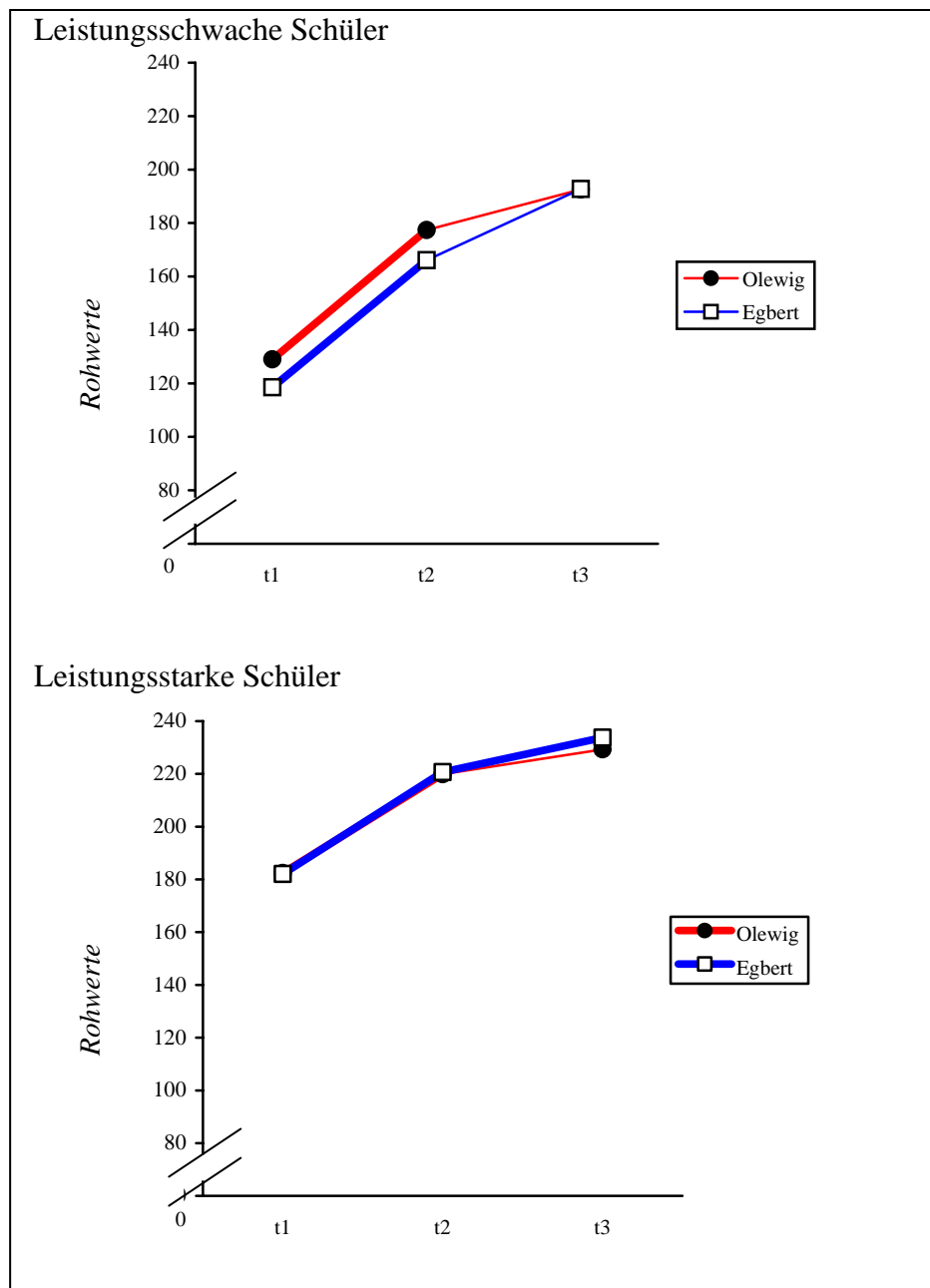


Abbildung 8.56 Absolute Rohwerte im DL-KG über die drei Messzeitpunkte (t_1 - t_3) für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 4 (theoretischer Range: 0 bis 240).

In der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* zeigt sich im Vergleich der beiden Schulen eine signifikant höhere Anzahl abgeschriebener Silben auf Seiten der Schülerinnen und Schüler der Experimentalschule (Olewig). Allerdings machen die Schülerinnen und Schüler aus der Experimentalschule (Olewig) auch signifikant mehr Fehler beim Abschreiben als die Schülerinnen und Schüler aus der Kontrollschule (Egbert). Schließlich erzielen die Schülerinnen und Schüler der Kontrollschule (Egbert) signifikant höhere Werte für rezeptive Aufmerksamkeit als die der Experimentalschule (Olewig). Da der TPK in Gruppe 4 lediglich einmal durchgeführt wurde, würde die Darstellung der Ergebnisse in Form von Abbildungen keinen zusätzlichen Informationsgewinn liefern.

8.4.3 Prüfung von Rechen- und Rechtschreibleistungen

In der *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* zeigen sich in keinem der beiden Maße der Rechenleistung bedeutsame Unterschiede zwischen den beiden Schulen.

In der *Hamburger Schreibprobe (HSP)* zeigt sich ein für beide Schulen vergleichbarer, statistisch bedeutsamer Leistungszuwachs über die Zeit. Der Vergleich der beiden Schulen zu den beiden Messzeitpunkten ergibt keine bedeutsamen Unterschiede (vgl. Abb. 8.57).

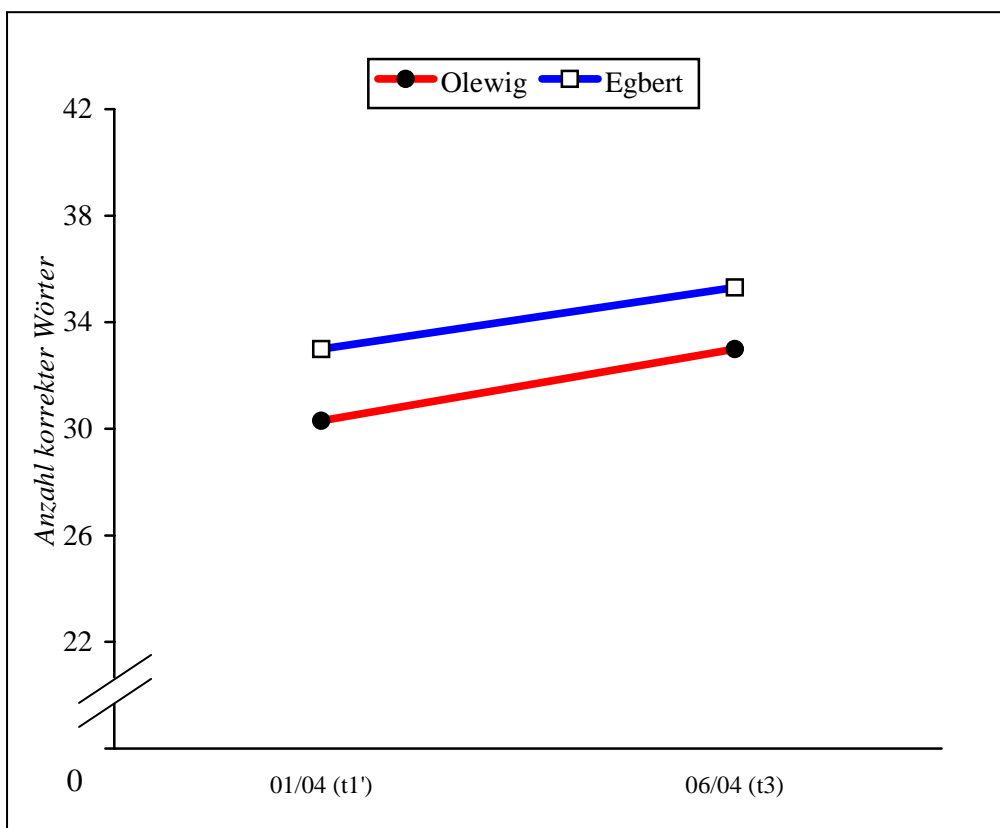


Abbildung 8.57 Anzahl korrekter Wörter im HSP über die beiden Messzeitpunkte (t_1' und t_3) für Gruppe 4 (theoretischer Range für t_1' und t_3 : 0 bis 42).

Die Erhebung zu t_1' wurde verzögert durchgeführt, um diese zeitgleich mit der Erhebung des ersten Schuljahres stattfinden zu lassen.

Für die leistungsschwachen Schülerinnen und Schüler beider Schulen lässt sich ein signifikanter Zuwachs der Leistung von t_1' nach t_3 nachweisen. Bei den leistungsstarken Schülerinnen und Schülern gilt dies nur für die Kontrollschule (Egbert, vgl. Abb. 8.58).

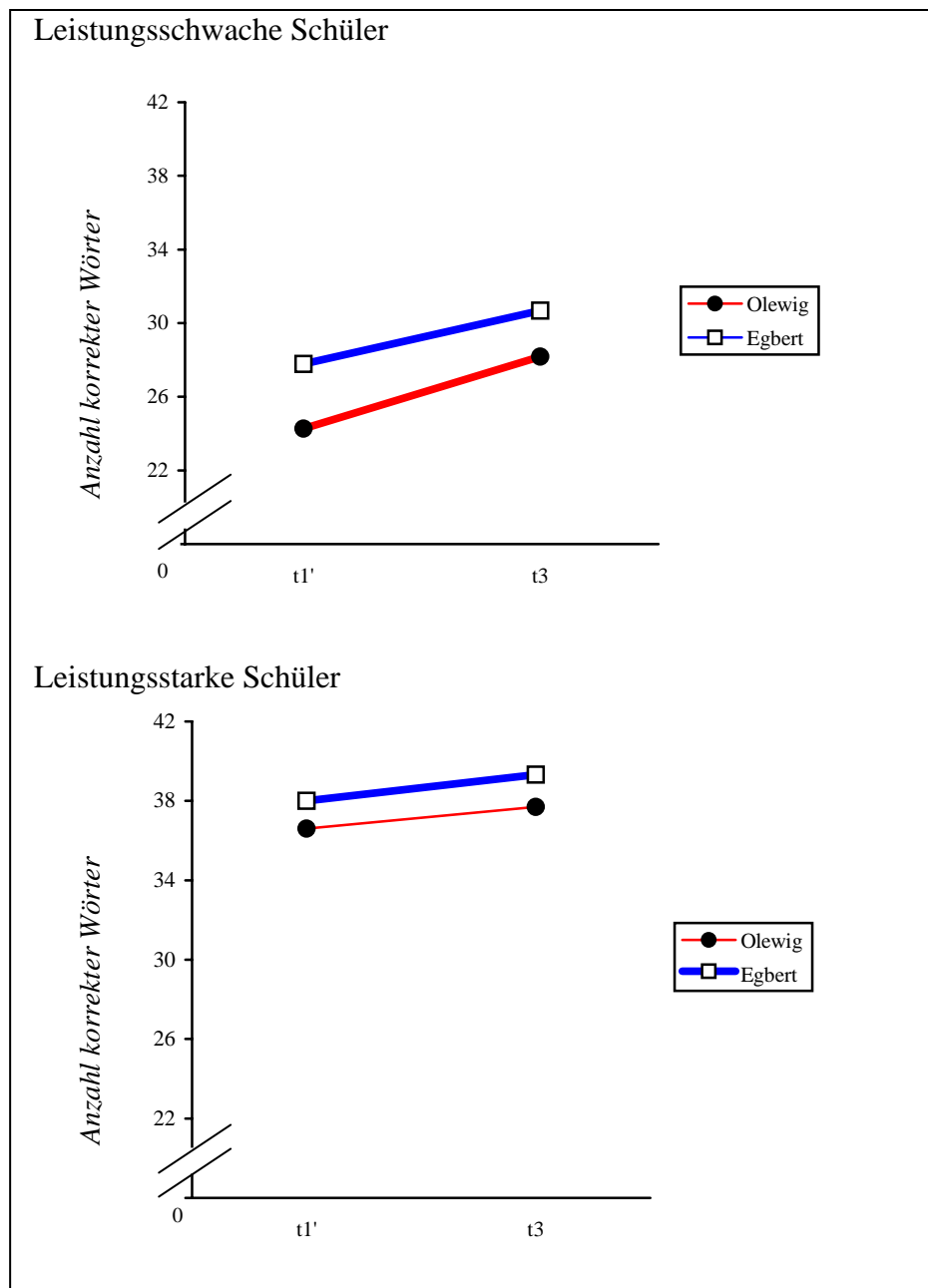


Abbildung 8.58 Anzahl korrekter Wörter im HSP über die beiden Messzeitpunkte (t_1' und t_3) für leistungsschwache resp. leistungsstarke Schüler aus Gruppe 4 (theoretischer Range für t_1' und t_3 : 0 bis 42). Die Erhebung zu t_1' wurde verzögert durchgeführt, um diese zeitgleich mit der Erhebung des ersten Schuljahres stattfinden zu lassen.

8.4.4 Aspekte der schulischen Integration

Im Fragebogen zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6) finden sich bedeutsame Unterschiede zwischen beiden Schulen auf der Dimension „Soziale Integration“, indem die Werte der Experimentalschule (Olewig) signifikant höher sind als die der Kontrollschule (Egbert, vgl. Abb. 8.59).

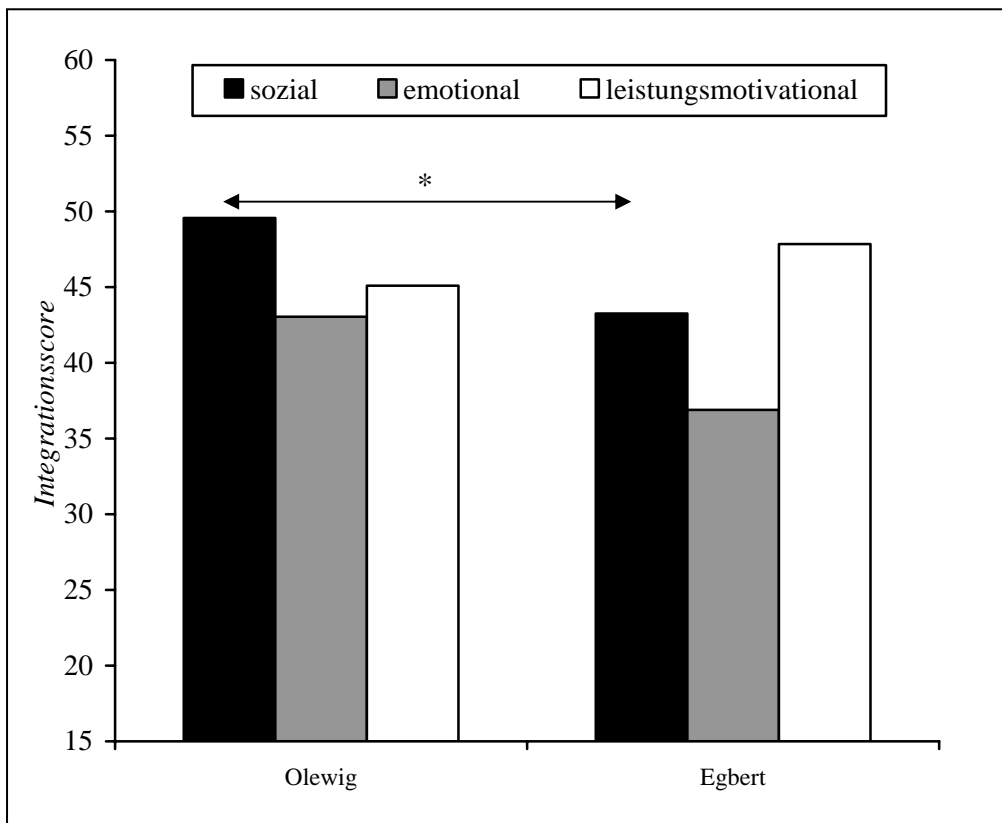


Abbildung 8.59 Scores für soziale, emotionale und leistungsmotivationale Integration in die Schule für Gruppe 4 (theoretischer Range: 15 bis 60). Die Werte für die schulische Integration der Schülerinnen und Schüler aus Gruppe 4 wurden nur einmal erhoben (t₃: 06/04).

Bei den schlecht integrierten Schülerinnen und Schülern liegt der Wert der Experimentalschule (Olewig) auf der Dimension „Emotionale Integration“ signifikant über dem Wert der Kontrollschule (Egbert). Auf der Dimension „Leistungsmotivationale Integration“ ist dagegen der Wert der Kontrollschule substantiell höher als der Wert der Experimentalschule. Die Schülerinnen und Schüler der Experimentalschule mit hohen Werten auf der jeweiligen Dimension weisen auf den Dimensionen „Soziale Integration“ und „Emotionale Integration“

signifikant höhere Integrationsscores auf als die Schülerinnen und Schüler aus der Kontrollschule (vgl Abb. 8.60).

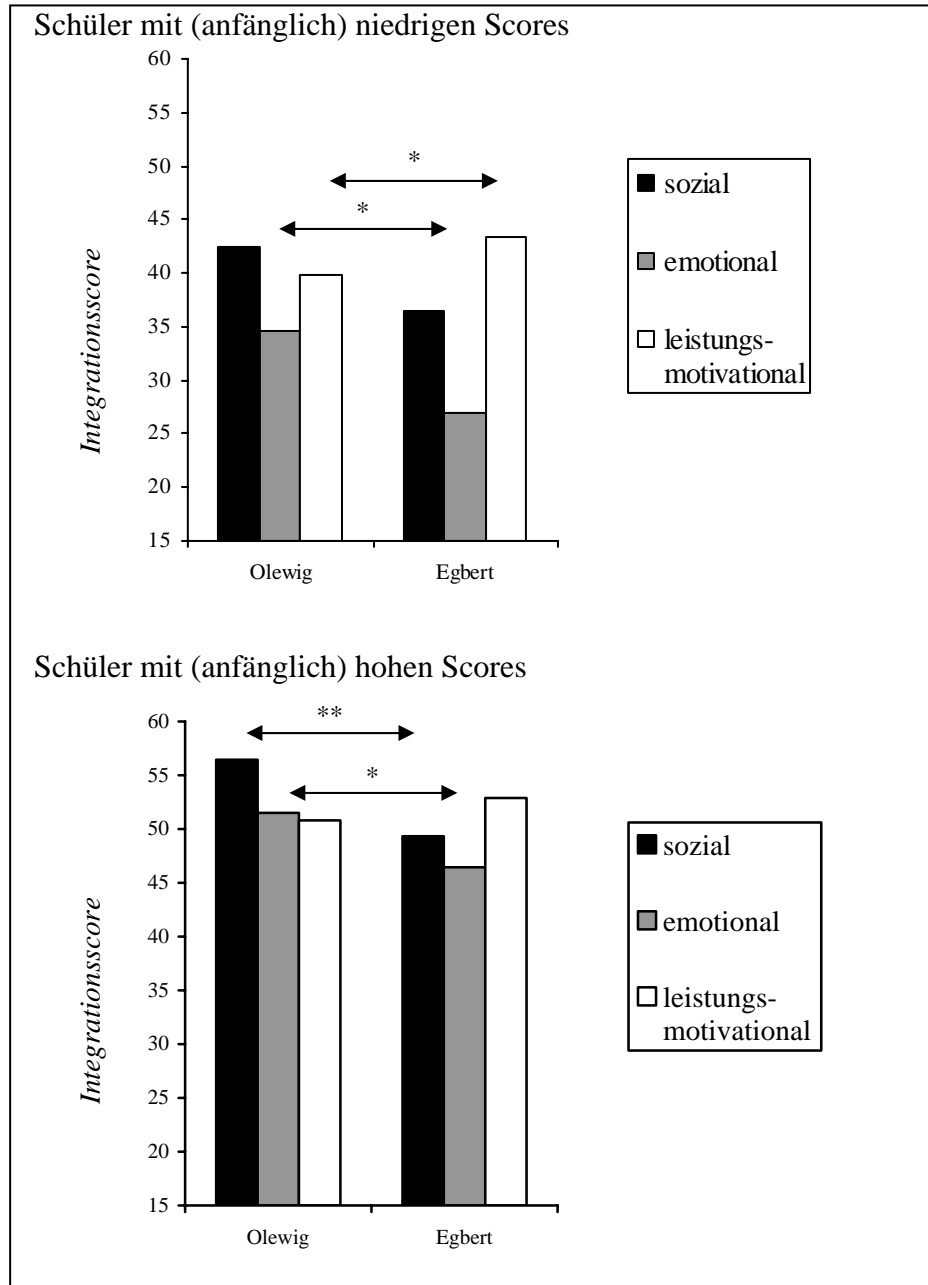


Abbildung 8.60 Scores für soziale, emotionale und leistungsmotivationale Integration in die Schule für zwei Teilstichproben (theoretischer Range: 15 bis 60). Die Werte für die schulische Integration der Schülerinnen und Schüler aus Gruppe 4 wurden nur einmal erhoben (t3: 06/04).

9 Zusammenfassende Betrachtung des Befundmaterials und abschließender Kommentar

In der Zusammenschau der vielen Detailbefunde, die in der vorangegangenen Ergebnisdarstellung dokumentiert wurden, lässt sich nochmals festhalten, was bereits einleitend gesagt worden war: Das Befundmuster ist uneinheitlich und stellt sich – in Abhängigkeit vom betrachteten Erhebungszeitpunkt, vom jeweils betrachteten Leistungsaspekt und in Abhängigkeit von der betrachteten Altersgruppe bzw. Schulklasse – höchst unterschiedlich dar. Gleichwohl ist eine Reihe von Befunden erwähnenswert, die auf eine partielle Überlegenheit der Olewiger Schülerinnen und Schüler in dem je betrachteten Aspekt hinweisen und die (auch) als positive Wirkung des Schachunterrichts gedeutet werden *können*. Im folgenden sei dies resümierend für die untersuchten Facetten der (intellektuellen) Leistungsfähigkeit dargestellt; abschließend wollen wir das Augenmerk lenken auf aus unserer Sicht interessante Unterschiede, die sich in den Aspekten der seitens der Kinder erlebten Integration in die Schule gezeigt haben.

Die relative Überlegenheit der Olewiger Kinder in einzelnen Leistungsmaßen ist unseren Befunden zufolge insbesondere in *Gruppe 2* zu beobachten, d.h. bei den Kindern, die bei Einführung des Schachunterrichts in Olewig und zu Beginn der Evaluationsstudie das zweite Schuljahr besucht hatten. In dieser Gruppe zeigt sich ein durchweg signifikanter Zuwachs in den Testleistungen der allgemeinen Intelligenz im ersten Jahr, der in ein signifikant höheres Leistungsniveau im Vergleich zu den Kindern aus der Grundschule Egbert mündet. Besonders zu erwähnen ist dabei, dass dieser Zuwachs im Niveau der allgemeinen Intelligenz, wie er in dieser Gruppe der Olewiger Kinder zu beobachten ist, insbesondere auf einen Zuwachs in den Intelligenztestleistungen bei den anfänglich als „*leistungsschwächer*“ eingestuften Kindern zurückgeht, während die bereits zu Beginn der Evaluationsstudie in diesem Bereich als „*leistungsstark*“ eingestuften Kinder keine signifikanten Verbesserungen über die Zeit und auch nicht im Vergleich zu den Kindern in Egbert aufweisen. Geht man noch weiter ins Detail, so ist zu erkennen, dass dieser Zuwachs in den Intelligenztestleistungen bei den anfänglich leistungsschwächeren Kindern insbesondere zurückzuführen ist auf eine signifikante Leistungsverbesserung in zwei Untertests, die produktive und reproduktive Aspekte der Wahrnehmung sowie visuelle Orientierung und Aufmerksamkeit abbilden. Fähigkeiten also, die gut mit dem

Schachspielen in Verbindung zu bringen sind. Was die Konzentrationsfähigkeit angeht, so findet man für diese Gruppe in beiden Schulen einen systematischen Anstieg über den Untersuchungszeitraum, besonders aber scheint die Konzentrationsfähigkeit bei den Schülerinnen und Schülern des zweiten Schuljahres in der Olewiger Schule zugenommen zu haben. Das gilt auch für ihre Rechenleistungen.

Hingegen fallen die Leistungsunterschiede zwischen Olewig und Egbert innerhalb der *Gruppe 1* (d.h. bei jenen, die zu Beginn des Schachunterrichts und der Evaluationsstudie in das Jahr eingeschult worden waren) nicht so deutlich ins Gewicht. Dies ist insofern bemerkenswert, als die Ergebnisse der VERA-Studie ja deutliche Unterschiede zu Gunsten der Kinder in Olewig nahelegen würden. Vielmehr gibt es im Niveau der allgemeinen Intelligenz wie auch in den Leistungen in einzelnen Intelligenzbereichen (d.h. in Untertests) leichte Unterschiede, die zu Gunsten der Kinder in Egbert ausfallen. Zudem lässt sich keine Zunahme in einzelnen Intelligenztestleistungen erkennen, die systematischer Natur wären. Wohl aber zeigt sich bei den Olewiger Kindern eine leichte Zunahme der Konzentrationsfähigkeit, die bei den Kindern in Egbert nicht in dem Maße deutlich wird; erkennbar wird bei den Olewiger Kindern dieser Gruppe (nicht bei jenen in Egbert) auch eine Zunahme in den Rechenleistungen und in den Schreibfertigkeiten, wobei diese Zunahme wiederum insbesondere auf eine solche bei den anfänglich leistungsschwächeren Kindern zurückgeht.

Für *Gruppe 3* zeigen sich kaum bedeutsame Unterschiede zwischen den Schulen mit Blick auf das Niveau der allgemeinen Intelligenz (Gesamttest), wohl aber lässt sich für einen Untertest („Klassifikation“, mit dessen Hilfe das beziehungsstiftende Denken der Schülerinnen und Schüler erfasst wird) im zweiten Jahr der Evaluationsstudie ein signifikanter Anstieg in der Experimentalschule nachweisen, der wiederum fast vollständig auf den Leistungsanstieg der anfänglich leistungsschwachen Schülerinnen und Schülern zurückgeführt werden kann. Mit Blick auf ein Maß der Konzentrationsfähigkeit (Gesamtzahl abgeschriebener Silben) zeigt sich eine Überlegenheit der Schülerinnen und Schüler in Olewig, die nicht nur in einem signifikanten Zuwachs zwischen dem zweiten und dritten Erhebungszeitpunkt liegt (das tut sie auch bei den Schülern von Egbert), sondern die auch in einen – im Vergleich zu Egbert – signifikant höheren Wert am Ende des zweiten Jahres der

Evaluationsstudie mündet. Die Tatsache, dass sich bei den Olewiger Kindern dieser Gruppe zugleich aber auch eine signifikant höhere Fehlerzahl als bei den Kindern aus Egbert zeigt, mag als Ausdruck einer höheren *Impulsivität* bei der Aufgabenbearbeitung durch die Olewiger Kinder gedeutet werden: höhere Leistungsmenge als ein Maß der Konzentrationsfähigkeit in Verbindung mit einer zugleich höheren Fehlerzahl. Gerade dies würde man von Kindern, die Schachunterricht erhalten, weniger erwarten als von Kindern, die nicht Schach spielen. Womöglich hat die hier beobachtete Impulsivität auch mit einem höheren Wohlbefinden in der Schule bei den Olewiger Kindern zu tun (siehe unten); jedenfalls sollte dieser Verhaltensaspekt in künftigen Studien noch stärker beachtet werden.

Schließlich zeigen sich für *Gruppe 4*, d.h. jenen Kindern, die bei Beginn des Schachunterrichts resp. der Evaluationsstudie bereits im vierten Schuljahr waren und die danach auf eine andere (weiterführende) Schule gewechselt sind, kaum bedeutsame Unterschiede zwischen den beiden Schulen, und zwar wieder im Niveau der allgemeinen Intelligenz noch in anderen Teilleistungen, die hier besondere Erwähnung verdienen würden. Ob der vermutete förderliche Einfluss des Schachspiels auf ausgewählte Facetten der geistigen Leistungsfähigkeit beschränkt bleibt auf *jüngere* Kinder und sich womöglich bei älteren Kindern nicht (mehr) in dem Maße nachweisen lässt, müssen weitere Studien zeigen. Unsere Daten sprechen jedenfalls in eine solche Richtung.

Dass das Schachspiel womöglich nicht (nur) die geistige Entwicklung fördern und die Konzentrationsfähigkeit steigern könnte, sondern sich auch positiv auf subjektive Facetten des Wohlbefindens auswirken könnte, wie sie hier mit Maßen der erlebten Integration in die Schule erfasst worden waren, ist eine mehr als plausible Ausgangsthese. Und in der Tat lassen sich die vorliegenden Ergebnisse als deutliche Belege für diese Ausgangsthese deuten: In *Gruppe 2* zeigte sich eine stabile emotionale Integration (d.h. hohes subjektives Wohlbefinden in der Schule) über den gesamten Untersuchungszeitraum bei den Kindern in Olewig, während diese in der bei den Kindern in Egbert sogar abnahm. Hier könnte das Schachspielen zu der beobachteten Stabilisierung beigetragen haben; denn wie (leider) nur unsystematische Beobachtungen nahelegen, hat das Schachspiel über diese drei Jahre einen immer breiteren Raum im Schulalltag eingenommen (z.B. Schachspielen in den Pausen, Schachspielen gegen „echte“ Schachmeister, Präsentationen auf

Schulfesten, dauerhafte Schachspielinstallationen, die jederzeit zugänglich waren etc.). All dies mag dem Wohlbefinden in der Schule dienlich gewesen sein. Auch in *Gruppe 3* traten bedeutsame Unterschiede zwischen den Kindern in der Grundschule Olewig und jenen in der Grundschule Egbert auf: Erstere fühlten sich in dem Erhebungszeitraum in leistungsmotivationaler Hinsicht besser integriert, d.h. sie zeigten ein höheres Vertrauen in ihre eigenen schulischen Fähigkeiten als die Kinder aus Egbert. Interessanterweise kommt dieser Unterschied wiederum vorwiegend dadurch zustande, dass die Kinder, die anfänglich niedrige Werte erreicht hatten („gering integrierte Kinder“), mehr und mehr an Selbstvertrauen hinzugewannen als jene, die bereits anfänglich ein vergleichsweise hohes schulisches Selbstvertrauen hatten. Schließlich ist in dieser Hinsicht noch ein statistisch bedeutsamer Unterschied für *Gruppe 4* zu erwähnen: Hier erreichten die Kinder aus Olewig im Maß der „sozialen Integration“ in die Schule deutlich höhere Werte, d.h. dass sie sich in der Schule subjektiv gut eingebunden und „verortet“ fühlten, und zwar in höherem Maße als die Kinder aus Egbert.

Abschließend müssen indes einige einschränkende Bedingungen genannt werden, die den Wert der Studie und ihrer Befunde partiell relativieren. Beide Schulen sollten sich gemäß der Zielsetzung der Evaluationsstudie darin unterscheiden, dass die Kinder der Grundschule Olewig regelmäßig und systematisch Schachunterricht erhielten, die Kinder der Grundschule Egbert hingegen nicht (laut Auskunft der Schulleitung). Indes stellte sich im Laufe der Evaluationsstudie heraus, dass es neben dem Schachunterricht auch freiwillige sog. Schach-Arbeitsgemeinschaften („Schach-AGs“) gab und dass dies sowohl in Olewig *wie in Egbert* der Fall war. Von den Egberter Kindern nahmen insgesamt allerdings nur sechs Kinder (davon zwei eher unregelmäßig), von den Olewiger Kindern insgesamt zwölf an diesen Schach-AGs teil. Insofern ist die Kontrastierung beider Schulen zwar nicht in dem Maße möglich, wie dies ursprünglich geplant und versuchsmethodisch erforderlich war. Auch wurde uns im Nachhinein berichtet, dass der Schachunterricht in der Grundschule Olewig zeitweise ausfallen musste und dass es dort einen Wechsel in dem das Schachspiel unterrichtenden Lehrpersonal gab. Folglich sind die Bedingungen, die unsere „Experimentalschule“ von der „Kontrollschule“ differenzieren sollten, nicht so systematisch unterschiedlich, wie man sich dies für eine „saubere“ Interventions- und Evaluationsstudie wünschen würde. Doch angesichts der genannten Zahlenverhältnisse und der Tatsache, dass es in Egbert

keinen systematischen Schachunterricht gab, lässt sich der Vergleich beider Schulen dennoch rechtfertigen. Ob und inwieweit die Tatsache, dass sich in einigen Testverfahren die Kinder aus Olewig und die Kinder aus Egbert nicht bedeutsam voneinander unterscheiden, durch eine diesbezügliche mangelnde Differenzierung der Schulen begründet sein könnte, lässt sich indes mit unseren Daten nicht beantworten.

Auch gilt es zu beachten, dass der Schachunterricht in der Grundschule Olewig ja nur eine Schulstunde pro Woche umfasste (auch wenn man bereits dies als einen wichtigen Schritt bewerten darf). Dennoch muss man argumentieren, dass angesichts der unermesslichen Fülle und Vielfalt von Faktoren, die (potentiell) Einfluss auf die geistige Entwicklung im Grundschulalter nehmen können, eine Stunde Schachunterricht sozusagen nur „ein Tropfen auf den heißen Stein“ sein kann. Vor diesem Hintergrund sind die Unterschiede, die die vorliegende Studie in ausgewählten Aspekten der Leistungsfähigkeit und erlebten Integration zwischen den Kindern mit und den Kindern ohne Schachunterricht ermittelt hat, durchaus bemerkenswert. Doch muss ein letzter Anflug von Skepsis diesen Bericht beenden: Wir wissen zu wenig darüber, ob sich beide Schulen – über die Einführung des Schachunterrichtes hinaus – nicht auch in *anderen* hier relevanten Merkmalen voneinander unterscheiden; und wir wissen auch nichts darüber, ob sich ggf. nicht auch die Kinder resp. ihre Herkunftsfamilien, die in diesen Schulen unterrichtet werden, in relevanten Merkmalen unterscheiden. So lange diese Informationsgrundlage nicht gegeben ist, mögen wir die Unterschiede, so weit sie zu Gunsten der Grundschule in Olewig sprechen, gerne im Sinne der förderlichen Wirkung des Schachunterrichts interpretieren. Einen letzten „Beweis“ dafür konnte die vorliegende Evaluationsstudie indes nicht liefern.

Anhang: Tabellarische Darstellung der Ergebnisse

A 1 Tabellarische Darstellung der zentralen Kennwerte für Gruppe 1

Tabelle 1.1. Mittelwerte und Standardabweichungen der relativierten Rohpunktwerte sowie Wertebereiche der absoluten Rohpunktwerte im CFT 1 (t1 bis t4) und CFT 20 (t5) in den beiden Schulen für alle fünf Messzeitpunkte (theoretischer Range: CFT 1: 0 bis 60 / CFT 20: 0 bis 46).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 14	10/03	1	0.55	0.13	18	48
	02/04	2	0.72	0.12	32	54
	06/04	3	0.78	0.11	36	56
	07/05	4	0.85	0.07	44	57
	07/06	5	0.50	0.14	8	34
Egbert (KG) n = 18	10/03	1	0.59	0.11	22	46
	02/04	2	0.71	0.11	30	53
	06/04	3	0.79	0.08	37	54
	07/05	4	0.90	0.05	48	58
	07/06	5	0.59	0.11	20	36

Tabelle 1.2. Mittelwerte und Standardabweichungen der relativierten Rohpunktwerte sowie Wertebereiche der absoluten Rohpunktwerte im DL-KE (t1 bis t3) und DL-KG (t4 und t5) in beiden Schulen für alle fünf Messzeitpunkte (theoretischer Range: DL-KE: 0 bis 98 / DL-KG: 0 bis 240).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 14	10/03	1	0.52	0.17	20	70
	02/04	2	0.59	0.19	26	84
	06/04	3	0.88	0.20	28	98
	07/05	4	0.58	0.14	56	188
	07/06	5	0.76	0.15	113	228
Egbert (KG) n = 18	10/03	1	0.50	0.18	21	85
	02/04	2	0.70	0.23	31	98
	06/04	3	0.77	0.21	37	98
	07/05	4	0.63	0.14	81	227
	07/06	5	0.73	0.19	69	240

Tabelle 1.3. Mittelwerte und Standardabweichungen der relativen Anzahl korrekt geschriebener Wörter sowie Werte-Range der absoluten Anzahl korrekt geschriebener Wörter im HSP in den beiden Schulen für alle vier Messzeitpunkte (theoretischer Range: t1 (Egbert): 0 bis 10; t1 (Olewig) und t2: 0 bis 14; t3: 0 bis 30; t4: 0 bis 38).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 16	01/04	1	0.50	0.23	1	12
	06/04	2	0.58	0.15	5	13
	07/05	3	0.72	0.16	13	30
	07/06	4	0.78	0.13	21	38
Egbert (KG) n = 19	01/04	1	0.47	0.22	0	8
	06/04	2	0.59	0.14	6	13
	07/05	3	0.69	0.15	12	29
	07/06	4	0.82	0.14	17	38

Tabelle 1.4. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Soziale Integration“ im FDI für alle drei Messzeitpunkte (theoretischer Range: 15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 13	06/04	1	47.31	10.94	27	59
	07/05	2	47.77	9.19	25	56
	07/06	3	47.69	7.15	27	55
Egbert (KG) n = 11	06/04	1	50.73	7.09	33	58
	07/05	2	44.73	10.88	20	58
	07/06	3	45.18	10.22	22	59

Tabelle 1.5. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Emotionale Integration“ im FDI für alle drei Messzeitpunkte (theoretischer Range: 15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 13	06/04	1	44.85	14.65	15	57
	07/05	2	37.38	16.14	15	59
	07/06	3	33.92	15.79	15	55
Egbert (KG) n = 11	06/04	1	51.18	7.48	37	59
	07/05	2	45.18	9.20	27	54
	07/06	3	39.82	11.50	21	56

Tabelle 1.6. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Leistungsmotivationale Integration“ im FDI für alle drei Messzeitpunkte (theoretischer Range:15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 13	06/04	1	43.85	10.63	21	54
	07/05	2	46.08	11.09	24	59
	07/06	3	41.00	11.92	19	57
Egbert (KG) n = 11	06/04	1	51.18	5.15	43	58
	07/05	2	47.36	7.74	36	60
	07/06	3	47.36	6.76	37	58

A 2 Tabellarische Darstellung der zentralen Kennwerte für Gruppe 2

Tabelle 2.1. Mittelwerte und Standardabweichungen der relativierten Rohpunktwerte sowie Werte-Range der absoluten Rohpunktwerte im CFT 1 (t1 bis t3) und CFT 20 (t4 und t5) in den beiden Schulen für alle fünf Messzeitpunkte (theoretischer Range: CFT 1: 0 bis 60 / CFT 20: 0 bis 46).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 18	10/03	1	0.76	0.07	37	53
	02/04	2	0.83	0.08	42	59
	06/04	3	0.90	0.06	44	59
	07/05	4	0.62	0.13	12	36
	07/06	5	0.69	0.10	23	40
Egbert (KG) n = 19	10/03	1	0.71	0.12	22	57
	02/04	2	0.76	0.09	36	57
	06/04	3	0.81	0.11	36	59
	07/05	4	0.63	0.13	14	43
	07/06	5	0.68	0.16	12	39

Tabelle 2.2. Mittelwerte, Standardabweichungen und Werte-Range der Rohpunktwerte im DL-KG in den beiden Schulen für alle fünf Messzeitpunkte (theoretischer Range: 0 bis 240).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 18	10/03	1	61.80	17.40	30	106
	02/04	2	116.73	40.30	60	185
	06/04	3	183.03	29.85	133	225
	07/05	4	211.00	23.91	173	240
	07/06	5	213.13	26.27	170	240
Egbert (KG) n = 18	10/03	1	76.04	31.57	16	126
	02/04	2	132.01	32.51	61	193
	06/04	3	170.39	31.54	122	235
	07/05	4	181.26	32.89	119	240
	07/06	5	195.08	33.79	97	240

Tabelle 2.3. Mittelwerte und Standardabweichungen der relativen Anzahl korrekt geschriebener Wörter sowie Werte-Range der absoluten Anzahl korrekt geschriebener Wörter im HSP in den beiden Schulen für alle vier Messzeitpunkte (theoretischer Range: t1: 0 bis 15; t2: 0 bis 30; t3: 0 bis 38; t4: 0 bis 42).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 18	01/04	1	0.68	0.16	6	15
	06/04	2	0.80	0.16	11	29
	07/05	3	0.87	0.14	16	38
	07/06	4	0.84	0.15	16	41
Egbert (KG) n = 15	01/04	1	0.68	0.18	6	15
	06/04	2	0.76	0.14	15	28
	07/05	3	0.84	0.11	23	38
	07/06	4	0.81	0.14	23	41

Tabelle 2.4. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Soziale Integration“ im FDI für alle drei Messzeitpunkte (theoretischer Range:15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 15	06/04	1	47.53	7.41	35	58
	07/05	2	48.47	6.29	36	55
	07/06	3	51.47	5.29	41	59
Egbert (KG) n = 16	06/04	1	47.06	6.28	35	56
	07/05	2	48.69	6.55	36	56
	07/06	3	49.38	7.02	30	58

Tabelle 2.5. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Emotionale Integration“ im FDI für alle drei Messzeitpunkte (theoretischer Range:15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 15	06/04	1	41.40	12.50	15	60
	07/05	2	41.27	13.99	15	60
	07/06	3	41.53	9.67	19	56
Egbert (KG) n = 16	06/04	1	47.49	10.61	24	59
	07/05	2	44.81	10.73	17	60
	07/06	3	38.50	12.02	21	60

Tabelle 2.6. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Leistungsmotivationale Integration“ im FDI für alle drei Messzeitpunkte (theoretischer Range:15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 15	06/04	1	41.40	8.54	23	53
	07/05	2	43.07	9.39	24	60
	07/06	3	49.47	5.34	40	60
Egbert (KG) n = 16	06/04	1	44.88	8.28	30	57
	07/05	2	43.69	8.62	32	59
	07/06	3	44.63	9.41	31	60

A 3 Tabellarische Darstellung der zentralen Kennwerte für Gruppe 3

Tabelle 3.1. Mittelwerte, Standardabweichungen und Werte-Range der Rohpunktwerte im CFT 20 in den beiden Schulen für alle vier Messzeitpunkte (theoretischer Range: 0 bis 46).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 15	10/03	1	26.33	4.48	13	32
	02/04	2	26.53	5.63	15	34
	06/04	3	29.13	5.59	14	36
	07/05	4	33.67	4.97	20	40
Egbert (KG) n = 19	10/03	1	21.74	10.03	6	35
	02/04	2	27.74	6.85	16	39
	06/04	3	30.11	5.49	19	39
	07/05	4	32.79	5.97	23	43

Tabelle 3.2. Mittelwerte, Standardabweichungen und Werte-Range der Rohpunktwerte im DL-KG in den beiden Schulen für alle vier Messzeitpunkte (theoretischer Range: 0 bis 240).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 15	10/03	1	163.06	35.05	98	211
	02/04	2	194.56	40.86	82	240
	06/04	3	207.57	30.48	141	240
	07/05	4	218.85	29.96	134	240
Egbert (KG) n = 19	10/03	1	152.15	40.71	85	236
	02/04	2	188.47	42.84	104	240
	06/04	3	210.76	27.77	141	240
	07/05	4	211.42	35.44	113	240

Tabelle 3.3. Mittelwerte und Standardabweichungen der relativen Anzahl korrekt geschriebener Wörter sowie Werte-Range der absoluten Anzahl korrekt geschriebener Wörter im HSP in den beiden Schulen für alle drei Messzeitpunkte (theoretischer Range t1 und t2: 0 bis 38 / t3: 0 bis 42).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 17	01/04	1	0.70	0.12	17	32
	06/04	2	0.77	0.09	24	33
	07/05	3	0.84	0.11	22	41
Egbert (KG) n = 18	01/04	1	0.71	0.15	17	36
	06/04	2	0.81	0.13	20	37
	07/05	3	0.83	0.16	17	42

Tabelle 3.4. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Soziale Integration“ im FDI (theoretischer Range:15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 17	06/04	1	45.71	7.80	22	56
	07/05	2	47.94	11.19	21	60
Egbert (KG) n = 20	06/04	1	45.45	8.76	24	59
	07/05	2	49.15	7.59	32	59

Tabelle 3.5. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Emotionale Integration“ im FDI (theoretischer Range:15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 17	06/04	1	38.94	7.73	17	52
	07/05	2	41.29	7.54	30	53
Egbert (KG) n = 20	06/04	1	42.75	11.99	21	60
	07/05	2	43.80	10.66	26	60

Tabelle 3.6. Mittelwerte, Standardabweichungen und Werte-Range auf der Dimension „Leistungsmotivationale Integration“ im FDI (theoretischer Range:15 bis 60).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 17	06/04	1	48.29	4.78	41	58
	07/05	2	50.41	5.67	39	59
Egbert (KG) n = 20	06/04	1	42.90	7.15	24	55
	07/05	2	45.50	8.11	33	60

A 4 Tabellarische Darstellung der zentralen Kennwerte für Gruppe 4

Tabelle 4.1. Mittelwerte, Standardabweichungen und Werte-Range der Rohpunktwerte im CFT 20 in den beiden Schulen für alle drei Messzeitpunkte (theoretischer Range: 0 bis 46)

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 20	10/03	1	27.25	4.73	11	35
	02/04	2	29.50	4.93	5	37
	06/04	3	32.20	4.51	7	41
Egbert (KG) n = 17	10/03	1	28.76	5.29	15	36
	02/04	2	30.88	6.37	14	38
	06/04	3	34.12	4.40	24	40

Tabelle 4.2. Mittelwerte, Standardabweichungen und Werte-Range der Rohpunktwerte im DL-KG in den beiden Schulen für alle drei Messzeitpunkte (theoretischer Range: 0 bis 240).

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 21	10/03	1	152.48	36.43	88	229
	02/04	2	194.77	36.29	117	240
	06/04	3	207.79	35.01	111	240
Egbert (KG) n = 17	10/03	1	150.08	39.74	54	207
	02/04	2	192.11	39.81	108	240
	06/04	3	211.77	31.78	146	240

Tabelle 4.3. Mittelwerte, Standardabweichungen und Werte-Range der Anzahl korrekt geschriebener Wörter im HSP in den beiden Schulen für beide Messzeitpunkte (theoretischer Range: t1 und t2: 0 bis 42)

Schule	Messzeitpunkt		Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 23	01/04	1	30.30	7.00	18	41
	06/04	2	33.00	6.72	18	41
Egbert (KG) n = 19	01/04	1	33.00	5.95	20	40
	06/04	2	35.31	5.68	21	42

Tabelle 4.4. Mittelwerte, Standardabweichungen und Werte-Range im FDI (theoretischer Range:15 bis 60)

Schule	Integration	Mittelwert	Standardabweichung	Minimalwert	Maximalwert
Olewig (EG) n = 23	sozial	49.57	8.60	29	59
	emotional	43.04	9.59	23	58
	leistungsmotivational	45.09	6.14	33	57
Egbert (KG) n = 19	sozial	43.26	8.98	15	59
	emotional	36.89	11.32	16	53
	leistungsmotivational	47.84	5.30	36	55