

Entwicklungsverlauf der Selbsteinschätzung
bei Kindern im Kindergarten und den
ersten beiden Schuljahren
Ein Längsschnitt

Development of Self-Concept of
Young children. A longitudinal Study

Inauguraldissertation der
Philosophisch–historischen Fakultät
der Universität Bern zur Erlangung
der Doktorwürde

vorgelegt von:

Reto Carlo Catani
von Bern und Monsano (It)

2004

Kontakt: r.c.catani@bluewin.ch

Von der philosophisch- historischen Fakultät auf Antrag von Prof. Dr. phil. A. Flammer und
Prof. Dr. phil. F.Oser angenommen.

Bern, den 5.November 2004. Der Dekan: Prof. Dr. Reinhard Schulze

Inhaltsverzeichnis

Vorwort

Abstract

1. Einleitung

1.1	Wie exakt schätzen Kinder ihre Leistungsfähigkeit ein?	9
1.2	Befunde zur Entwicklung der Höhe der Selbsteinschätzung	12
1.3	Probleme mit dem diskontinuierlichen Verlauf	14
1.4	Die Fragestellungen dieser Arbeit	15
1.5	Zum Inhalt und Aufbau des Theorieteils	16

2. Die Entwicklung der Selbsteinschätzung im Schuleingangsalter

2.1	Einleitung	17
2.2	Determinanten der Selbsteinschätzung	18
2.2.1	Die Determinanten der Selbsteinschätzung im Überblick	18
2.2.2	Determinanten des gesellschaftlich/ schulischen Kontextes	19
	2.2.2.1 Die steigende Bedeutung der Selbstverwirklichung	
	2.2.2.2 Der normative Kontext: Die Förderung eines positiven Selbst- konzeptes in der Schule, dargestellt am Beispiel des Kantons Bern	
	2.2.2.3 Selbstgesteuerte Unterrichtsformen setzen Selbstvertrauen voraus	
	2.2.2.4 Beurteilung und Schwierigkeitsgrad von Lernaufgaben	
2.2.3	Determinanten der Schulklasse	31
2.2.4	Determinanten des Individuums	32
	2.2.4.1 Bisherige Leistung und persönlicher Standard	
	2.2.4.2 Der Wandel in der Bezugsnormorientierung	
	2.2.4.3 Der Einfluss der Lernfreude	
	2.2.4.4 Der Aufbau und die Ausdifferenzierung des Attributionskonzeptes	
2.2.5	Zusammenfassung	41

2.3	Befunde zur Entwicklung und Ausdifferenzierung spezifischer Kausalitätskonzepte	43
2.3.1	Anstrengung und Fähigkeit als Ursache von Erfolg unterscheiden lernen	43
2.3.2	Die Beurteilung der Aufgabenschwierigkeit	44
2.3.3	Wirkungen der Ausdifferenzierung von Fähigkeit und Anstrengung auf Höhe und Genauigkeit der Selbsteinschätzung	46
2.4	Empirische Befunde zur Verlaufsgestalt der Selbsteinschätzung im Kindergarten- und Grundschulalter	47
2.5	Unterschiedliche Einschätzhöhe bei Knaben und Mädchen	55
2.6.	Fragestellungen und Hypothesen	59
2.6.1	Verlaufsgestalt der Selbsteinschätzung	59
2.6.2	Die Schätzgenauigkeit nimmt zu	62
2.6.3	Geschlechtsunterschiede	63
2.6.4	Unterschiedliche Sozialisierungsbedingungen von Stadt- und Landkindern	64

3. Methode

3.1	Die Stichprobe der Längsschnittuntersuchung	66
3.2	Der Vergleich von Längsschnitt- und Kontrollgruppen	69
3.3	Vorgehen	73
3.4	Erhebungsinstrumente	76
3.4.1	Leistungstests in Mathematik	76
3.4.2	Die Instrumente für den Leistungsbereich Sport	79
3.4.3	Die Operationalisierung der Selbsteinschätzung	83
3.5	Der Untersuchungsplan	89

4. Darstellung der Befunde und Interpretation

4.1	Übersicht	90
4.2	Ergebnisse in Mathematik	90
4.2.1	Höhe der Einschätzung im Schuleingangsalter	91
4.2.2	Befunde zur Einschätzhöhe unter Berücksichtigung der Variablen Klassenstufe, Geschlecht und Milieu	93
4.2.3	Geschlechts- und Milieuunterschiede	95
4.2.4	Mittelwerte der Einschätzhöhe unter Berücksichtigung der Leistung und der unabhängigen Variablen Geschlecht und Milieu	96

4.2.5	Die Leistungen in Mathematik	98
4.2.6	Über- oder Unterschätzung als Differenz zwischen Einschätzung und Leistung in Treppenstufen	102
4.2.7	Überprüfung der Zusammenhangshypothesen	106
4.2.8	Häufigkeitsverteilungen und Betrachtung von Extremeinschätzungen	109
4.3	Ergebnisse im Sport: Monopedales Überhüpfen	114
4.3.1	Die Leistungen in monopedaalem Überhüpfen	115
4.3.2	Höhe der Einschätzung auf der Treppenskala	117
4.3.3	Befunde zur Einschätzhöhe unter Berücksichtigung von Klassenstufe, Geschlecht und Milieu	119
4.3.4	Über- oder Unterschätzung im monopedaalen Überhüpfen	123
4.3.5	Überprüfung der Zusammenhangshypothesen	125
4.4	Ergebnisse im Sport: Balancieren rückwärts	130
4.4.1	Die Leistungen in Balancieren rückwärts	130
4.4.2	Höhe der Einschätzung auf der Treppenskala	132
4.4.3	Befunde zur Einschätzhöhe unter Berücksichtigung von Klassenstufe, Geschlecht und Milieu	134
4.4.4	Über- oder Unterschätzung als Differenz zwischen Einschätzung und Leistung im Balancieren rückwärts	139
4.4.5	Überprüfung der Zusammenhangshypothesen	144
4.5	Schlussbetrachtung und Ausblick	147
4.5.1	Höhe und Verlauf der Selbsteinschätzung	147
4.5.2	Geschlechts- und Milieudifferenzen	150
4.5.3	Methodisches Vorgehen und Instrumente	152
5.	Literaturverzeichnis	153
6.	Anhang	161

Vorwort

„Spieglein, Spieglein an der Wand, wer ist die Schönste im ganzen Land?“

Mit diesem Spruch befragt die Königin im Grimm'schen Märchen vom „Schneewittchen“ ihren Spiegel, um herauszufinden, wer denn die schönste Frau im Lande sei. Die Geschichte endet für die Königin tragisch: Weil sie es nicht ertragen kann, nur die zweitschönste Frau zu sein, stürzt sie sich und andere ins Unglück. Die Königin möchte nicht einfach schön sein, nein sie ist nur glücklich, wenn keine Frau schöner ist als sie und diese „Hoffärtigkeit“, so die Moral der Geschichte, bezahlt sie mit dem Leben.

Glück und Zufriedenheit mit sich Selbst sind auch heute in vielfältiger Weise damit verknüpft, wie gut man in bestimmten Dingen ist oder wie gut man etwas kann, Wettbewerbssituationen gehören zum persönlichen und medialen Alltag unserer Kultur. Man misst sich häufig, so wie die Königin, vorwiegend an den Leistungen anderer um zu beurteilen, ob etwas gut gelungen ist oder nicht. Demgegenüber können wir unsere Zufriedenheit ja auch daraus schöpfen, wie viele Fortschritte wir gemacht haben: Luca, mein achtjähriger Sohn, der zur Zeit die zweite Klasse besucht, kann stundenlang erzählen, was er im Fussballtraining gerade dazulernt und wie gut er doch nun allgemein im Fussball geworden ist. Er orientiert sich bei der Selbstbeurteilung noch vorwiegend am eigenen Lernfortschritt, so wie dies die Kinder seiner Altersgruppe in der vorliegenden Arbeit auch tun: Wenn sie etwas dazu gelernt haben, dann sind sie nicht nur besser, sondern eben auch grundsätzlich „gut“, sprich leistungsstark. - Damit sind wir mitten im Thema der vorliegenden Arbeit. Es geht um die Frage, wie sich Kinder im Schuleingangsalter, genauer vom Kindergarten bis zur zweiten Klasse, in zentralen Leistungsbereichen selber einschätzen. Neigen sie dazu, sich im Vergleich zu den anderen Kindern zu hoch einzuschätzen? Wird die Einschätzung mit den Jahren durchschnittlich realistischer, sinkt sie oder nimmt die Einschätzungshöhe noch zu? Schätzen sich Knaben generell anders ein als Mädchen? Unterscheiden sich die Einschätzungshöhen je danach, ob es sich um Kinder aus städtischen oder eher ländlichen Gebieten handelt?

Diese Fragen sind das Thema dieser Arbeit, sie werden möglichst präzise beantwortet. Das wäre allerdings nicht möglich, wenn die untersuchten Kinder nicht mit viel Interesse und Engagement bereit gewesen wären, sich einer jährlich wiederkehrenden, etwa halbstündigen Befragung und Prüfung zu unterziehen. Deshalb gilt mein erster Dank den fast 450 Kindern aus dem Berner Oberland, dem Emmen- und Gürbetal, sowie dem Schwarzenbur-

gerland, die an den Untersuchungen teilgenommen haben. Dass dies überhaupt möglich wurde, verdanke ich zudem den spontanen und zuverlässigen Kindergärtnerinnen und Lehrerinnen, die die Untersuchungen jeweils in ihren Alltag einplanten und einen möglichst reibungslosen Ablauf ermöglichten: Ihr Interesse an der Fragestellung war hoch und alle wirkten gerne an diesem Projekt mit.

Selbstverständlich kann eine solche Untersuchung nur gelingen, wenn sich die Fachwissenschaft dafür einsetzt. Besonderer Dank gilt deshalb Prof. Dr. August Flammer, der sich, trotz meines fortgeschrittenen Alters, vorstellen konnte, mich bei der Arbeit zu betreuen. Mit seinem differenzierten Wissen hat er viel zum Gelingen dieser Arbeit beigetragen und mit seiner Mischung aus Gewährenlassen und hartnäckigem Fragen einerseits die Motivation des Autors erhalten, dann aber auch dessen Denken geschärft. Prof. Dr. Françoise Alsaker und Prof. Dr. Fritz Oser möchte ich ebenfalls für ihren Beitrag zum Gelingen der Arbeit danken. Ein weiterer Dank geht an Prof. Dr. Alex Grob und meinen ehemaligen Kollegen Dr. Markus Kübler, die mich zu Beginn der Arbeit ermunterten und unterstützten.

David Schmid, Stefan Valkanover und Christof Nägele haben mich durch Fachgespräche in meiner Arbeit mehrmals unterstützt. Nathalie Bider, Daniel Schütz, David Schmid und Christina Zingg haben bei der Datenerhebung ebenso mitgeholfen, wie die Studierenden der LAB-Kurse 6 und 9 des Seminars Spiez: Ohne ihre zuverlässige Arbeit wäre es nicht möglich gewesen die Daten in der benötigten Qualität zu erheben, besten Dank! Gedankt sei auch meinem ehemaligen Vorgesetzten in Spiez und den involvierten Schulbehörden im Berner Oberland, die meine Arbeit mit Wohlwollen unterstützten.

„Last but not least“ gebührt mein Dank meiner Familie, insbesondere meiner Frau, die mir über Jahre meine Stunden im „stillen Kämmerlein“ zugestanden hat, auch wenn anderes zu tun gewesen wäre.

Abstract

Because of the small number of longitudinal studies, little is known about the development and accuracy of the mean levels of competence self-ratings of younger children. Instruments and scales are rare because young children have to be interviewed orally, not by paper-pencil-tests. In a longitudinal study we examined the self-ratings in the subjects „mathematics“ and „sports“. A total of 200 children were assessed in kindergarten (age 6.8 years), first and second grade. The children were interviewed after 8 months of instruction in the same classes, in May / June 1999, 2000 and 2001. The change from kindergarten to school brought only few change in the compositions of the classes. About 80% of the second graders also visited together the same kindergarden. The sample was taken in the Canton Berne, Switzerland and represents urban and rural children and in equal proportions boys and girls.

Each child was interviewed individually by students of the Teachers-Training-College Spiez, Canton Berne. First the child had to rate on which level its performance in mathematics and in two sport tasks, „Balancing Backwards“ and „High Jumping“, will be. The rating scale included 17 levels and was presented as a staircase. On the steps children of the same age were depicted. On the top steps stood the children attending the next grade, on the down steps the children representing the lower grade children. After the rating, the children made a performance test in mathematics and both sport tasks. Because of possible effects of repeated measuring, the results of the examined children were compared with data from control groups (first grade: $n=101$; second grade: $n=98$). The ratings differed not significantly.

Surprisingly the self-ratings in mathematics raised significantly, from $M = 0.50$ in Kindergarten to $M=1.62$ in first grade, to $M=1.16$ in second grade (scale= -8 to +8). Most of the newer studies found declining self-ratings (Helmke, 1991; Marsh, Craven, Debus, 1998). The self-ratings in „High-Jumping“ stayed stable: Kindergarten: $M=0.57$, first grade: $M=0.73$, second grade: $M=0.80$. The differences were not significant in any combination. In „Balancing-Backwards“ the means did not change significantly either: Kindergarten: $M=0.56$, first grade: $M=0.19$, second grade $M=0.18$. The often reported declining of the self-rated performance can not be confirmed by our studies. After two years of schooling our children have still a high self-confidence in mathematics and sports.

The accuracy (self-rating x performance) in mathematics started with $r=0.289^{**}$, to $r=0.181^*$, to $r=0.348^{**}$ and was higher than in all compareable studies and significant in each year ($*p<0.05$; $**p<0.01$). In sports the accuracy in Kindergarten was $r= -0.031$ (jumping) and

$r=0.037$ (balancing), in first grade $r=0.183^*$ and $r=0.131$, in second grade $r=0.103$ and $r=0.193^{**}$. In sports the children rated their performance level not only with the help of the staircase-scale, but also on the task: They estimated, how high they will jump and how far they will be able to balance without producing a mistake. The accuracy (self-rating on task \times performance) in Kindergarten was $r=0.141^*$ (jumping), $r=0.140^*$ (balancing), in first grade $r=0.238^{**}$ and $r=0.132$, in second grade 0.312^{**} and $r=0.182^*$. Our findings, related to accuracy show higher correlations than older studies do (for mathematics). For sports they are about at the same level as reported by Valkanover (2003) and Ingold (2002).

With ANOVAs we tested cohort (rural versus urban classes) and gender differences. Only few interactions were found: In balancing the development of overestimation was significant different for boys and girls. No significant differences in self-ratings were found in all studies between rural and urban children. The boys self-ratings in mathematics and jumping were significantly higher than the girls ratings. In balancing no significant difference was found between boys and girls. The performance of boys and girls in some tasks was not, as expected, equal, sometimes the boys were significantly better. With ANCOVAs we controlled the influence of performance. Under these conditions, the boys self-ratings did not differ significantly from the girls self-ratings in mathematics and in sports (one exception: in jumping, second grade, the boys rated significantly higher than girls did). It is important to control the performance before statements on gender differences are made: A lot of older studies did not control the performance or it was controlled by measuring it indirectly, they just asked the parents or teachers.

Our results on the gender differences correspond with the findings of Valeski and Stipek (2001), Cole et al. (1999), and Marsh et al. (1989). Demény (2002) found significant gender differences on second grade (performance not controlled). In Helmke's studies (1997), the second grade boys only tendentially rated their performance in mathematics higher than girls did.

1. Einleitung

Mit Fragen zur Entwicklung der Selbsteinschätzung in relevanten Leistungsbereichen hat sich die psychologische Forschung über Jahrzehnte beschäftigt. Entsprechend sind die Befunde gut belegt und es herrscht beispielsweise Konsens darüber, dass sich Menschen im Vergleich zu anderen Menschen durchschnittlich leistungsmässig überschätzen. Auch die Entwicklung der Höhe der Selbsteinschätzung wurde eingehend untersucht. Es ist gut belegt, dass die Einschätzhöhe in schulrelevanten Leistungsbereichen über die Schuljahre hinweg abnimmt. Wenige Studien gibt es dann zur Einschätzungsfähigkeit bei jüngeren Kindern, denn in diesem Alter ist die Datenaufnahme sehr aufwändig. Entsprechend klein sind dann manchmal die Stichproben. Im folgenden wollen wir uns ganz diesen jüngeren Kindern widmen, beginnend zum Zeitpunkt, in dem sie mit dem regelmässigen Lernen in grösseren Gruppen erste Erfahrungen sammeln, im Kindergartenalter also. Ein und zwei Jahre später werden die gleichen Kinder nochmals untersucht. In den folgenden drei Unterkapiteln werden die untenstehenden Fragen bearbeitet:

- Wie realistisch schätzen sich Kinder im Kindergarten und den ersten Schuljahren ein?
- Werden Kinder in ihren Einschätzungen vom Kindergarten bis zur zweiten Klasse genauer?
- Stimmt die Vermutung, dass Unterricht tatsächlich zu einer Abnahme der Höhe der Selbsteinschätzung führt?

1.1 Wie exakt schätzen Kinder ihre Leistungsfähigkeit ein?

Das Lernen in Kindergarten und Schule findet in einem öffentlichen, für alle teilnehmenden Personen sichtbaren Rahmen statt. Wie erleben eigentlich Kinder diese Lernsituationen? Wie leben die Schwächeren damit, dass andere fast immer besser sind als sie? Nehmen sie das überhaupt wahr oder orientieren sie sich bezüglich ihrer Leistungsfähigkeit an anderen Kriterien, etwa am Zuspruch von Lehrpersonen oder Eltern? Vom ersten Schultag an müssen Kinder vor Mitschülerinnen und -schülern singen, rennen, rechnen, zeichnen, lesen, usw. Bemerkten sie in den ersten Jahren überhaupt, dass andere dabei schneller, geschickter, tüchtiger sind? Empfinden Kinder dieses Lernen unter Wettbewerbsbedingungen auch so belastend, wie dies Erwachsene tun?

Lernen unter Wettbewerbsbedingungen ist nicht unbedenklich: Mietzel (1993) sieht insbesondere die leistungsschwächeren Kinder benachteiligt, sofern sie häufig und andauernd zu den Verlierern des Wettbewerbs gehören. Die Wettbewerbsverlierer neigen stark dazu, Gründe wie Zufall oder ihre Fähigkeiten als Ursache des eigenen Misserfolgs zu be-

zeichnen (Ames, 1978; Ames et al., 1977; Ames und Felker, 1979, alle zit.n.: Mietzel, 1993). Diese Ursachen entziehen sich der Beeinflussung durch das Kind, es reagiert darauf häufig mit Hilflosigkeit oder Ablehnung.

Kinder im Schuleingangsalter sind noch ganz am Anfang ihrer langen Schullaufbahn. Sie sind zum ersten Mal in ihrem Leben mit einer grossen Gruppe Gleichaltriger zusammen um den Lernalltag zu erfahren. Das Schwierige dieser Situation ist, dass sich die Kinder diesem Lernsetting nicht entziehen können, sind sie leistungsstark, dann werden sie auch von den anderen Kindern gelobt und sind entsprechend beliebt (Meyer, 1984). Da die Kinder einer Klasse alleine aufgrund ihres Alters zusammen den Kindergarten besuchen, werden sie alle mit den gleichen curricularen Anforderungen konfrontiert, unabhängig davon, ob sie diesen schon gewachsen sind, ja es gilt, allfällige Defizite möglichst früh zu erfassen. Kinder können also schon früh erfahren, dass sie oder Mitschülerinnen und –schüler leistungsschwächer sind als die meisten Kinder ihrer Klasse.

Schulartiges Lernen beginnt bereits im Kindergarten: Man lernt gemeinsam Lieder, denkt über Geschichten nach, stellt Gegenstände her, zeichnet Blumen, Häuser, Menschen und spielt im Sportunterricht in Mannschaften oder rennt im Stafettenlauf gegeneinander. Kinder lernen sich dadurch gegenseitig in ihren Stärken und Schwächen kennen, auch wenn die wichtigen Fertigkeiten wie das Lesen und Schreiben können oder das Rechnen noch nicht zum Lernpensum des Kindergartens gehören. Kinder machen also viele wichtige Leistungserfahrungen in sozialen Gruppen, denen sie sich nicht entziehen können. Die Frage ist nun, ob sie in ihrem Lerneifer und in ihrem Leistungsselbstbild durch die Gruppe beeinflusst werden.

An welchen Massstäben messen Kinder im Schuleingangsalter ihren Lernerfolg? Können sie sich überhaupt präzise einschätzen? Befunde von Nicholls (1978, 1979) zeigen, dass Kinder ihren Rangplatz innerhalb der eigenen Schulklasse erst im Alter von 9 bis 10 Jahren einigermaßen realistisch einschätzen. In der Untersuchung von Nicholls hatten die Kinder (N=16 pro Jahrgang) auf einer 30-stufigen Skala zu beurteilen, wie gut sie im Vergleich zu ihren Klassenkameraden lesen können. Ihre Selbsteinschätzung verglich Nicholls mit den Beurteilungen der Lehrkräfte (Tab.1-1-1):

Tabelle 1-1-1: Korrelation von subjektiver und objektiver Rangplatzschätzung bezüglich Lesefähigkeit (Nicholls, 1978)

Klasse	1.Kl.	2.Kl.	3.Kl.	4.Kl.
r	0.21	0.27	0.58*	0.71*

Die Genauigkeit der Selbsteinschätzung (accuracy), hier operationalisiert als Korrelation zwischen Rangplatzschätzung der Lehrperson und Einschätzung des Rangplatzes durch das Kind, erreichte erst in der dritten Klasse einen signifikanten Wert auf dem 5% -Niveau (für den Kindergarten liegen in dieser Untersuchung keine Einschätzungen durch die Kindergärtnerinnen vor). Mit der Genauigkeit der Selbsteinschätzung im vorliegenden Fall wird die Fähigkeit des Kindes gemessen, sich auf einer Skala ungefähr am richtigen Ort einzuschätzen. Würden Rangschätzung durch Kind und Lehrperson übereinstimmen, wäre die Korrelation perfekt. Der Verlauf in Tab.1-1-1 zeigt, dass in der dritten Klasse ein merklicher Anstieg erfolgt, wie entwickelt sich die Schätzgenauigkeit wohl vom Kindergarten zur ersten Klasse?

Die Schätzgenauigkeit wurde in der obigen Untersuchung als Korrelation von Selbsteinschätzung durch die Kinder und Fremdurteil durch Erziehungspersonen operationalisiert, es gibt aber noch andere Möglichkeiten: Als zweites mögliches Fremdurteil kommt auch die Leistungsbeurteilung in Form von Zeugnisnoten in Frage. Die dritte Möglichkeit besteht darin, die Leistungen der Kinder mit einem validen, reliablen Leistungstest zu messen. In diesem Fall wäre es dann möglich, die Leistungseinschätzung der Kinder (in Mathematik bin ich „genügend“, „gut“, „sehr gut“, usw.) mit den Messresultaten zu vergleichen. Allerdings ist ein Vergleich zwischen einer absoluten Bezugsnorm und der Selbsteinschätzung des Kindes unter Umständen problematisch: Reliable Leistungstests beinhalten meist Aufgaben, die nicht an einen bestimmten Lehrplan oder ein Lehrmittel gebunden sind. Je stärker das dem Unterricht zu Grunde gelegte Lehrmittel auf andere Lernziele baut als der Leistungstest, desto ungenauer misst der Leistungstest das, was das Kind gelernt hat, die Validität ist dann eingeschränkt. Eine Untersuchung dieser Art strebt deshalb eine möglichst hohe Validität des Messinstrumentes an, damit sich Kinder unter mathematischer Leistungsfähigkeit möglichst deckungsgleich dasselbe vorstellen, was der Leistungstest misst.

In Untersuchungen der Schuleingangsstufe kommt ein weiteres Problem dazu (beispielsweise bei Helmke, 1991). Im Kindergarten werden keine schriftlichen Leistungsbeurteilungen vorgenommen, am Ende der zweiten Klasse erhalten die Kinder dann Zeugnisse mit Noten (Ziffern, die die Leistung qualifizieren). Im Übergang vom Kindergarten zur ersten Klasse oder auch später wechselt deshalb in Untersuchungen häufig die Art der Fremdbeurteilung von einer Rangplatzeinschätzung zu Leistungsnoten oder sogar Leistungstests, das Aussenkriterium (Fremdurteil) ist nicht konstant (Tab.1-1-2):

Tabelle 1-1-2: Korrelation von Selbst- und Fremdbeurteilung in Mathematik (Helmke, 1991)

Klasse.	KG	1.Kl.	2.Kl.	3.Kl.
Rangreihe in Math.	0.15	0.16		
Note in Mathematik			0.32	0.34

Leider wechselte hier das Aussenkriterium nach der ersten Klasse von einer nicht auf Leistungstests beruhenden Beurteilung der Leistung (Erfahrungsbeurteilung durch die Lehrkraft) zu einer Beurteilung, welche aufgrund von Prüfungen im Fachbereich erfolgte. Da die Zunahme der Korrelation vom ersten zum zweiten Schuljahr im Vergleich zu den nachfolgenden Jahren bei Helmke doch ziemlich sprunghaft war, kann es durchaus sein, dass bei Anwendung entsprechender Leistungstests im Kindergarten und dem ersten Schuljahr höhere Korrelationen gefunden werden. Die Konstanthaltung des Aussenkriteriums wäre deshalb günstig, eignen würde sich etwa ein adaptiver Schulleistungstest, der bereits im Kleinkindalter eingesetzt werden kann.

1.2 Befunde zur Entwicklung der Höhe der Selbsteinschätzung

Am Ende der ersten Klasse können Kinder lesen und sind zunehmend in der Lage, jemandem etwas schriftlich mitzuteilen. Das ist für die meisten Menschen ein bedeutsamer Fortschritt. Ähnliches geschieht im Mathematikunterricht: Kinder können am Ende der ersten Klasse weit über hundert hinaus zählen, erkennen die Ziffern und können einfache Additionen und Subtraktionen ausführen. Ich vermute, dass ein vergleichbares Gefühl von Kompetenzzuwachs, wie es in diesen Fachbereichen im ersten Schuljahr erlebt wird, in der Volksschule erst wesentlich später wieder erfahren wird, beispielsweise bei der Einführung des Fremdsprachenunterrichts im fünften Schuljahr. Wenn nun die überwiegende Bezugsnorm im Schuleingangsalter der eigene Lernfortschritt ist, dann müsste sich das Kind am Ende der ersten Klasse als sehr kompetent einschätzen. Dies bestätigen Befunde: Der Beurteilung des eigenen Lernerfolges legen die Kinder zu Beginn der Schuleingangsphase den erfahrenen Lernfortschritt zu Grunde, erst Ende der zweiten Klasse wird die soziale Bezugsnorm bedeutsam (Nicholls, 1978; Ruble, Boggiano, Feldman & Loebel, 1980). Es verwundert deshalb nicht, dass die Einschätzung der eigenen Leistungsfähigkeit im ersten Schuljahr am höchsten ist und dann allmählich abnimmt. Auch dieser Befund konnte durch viele Untersuchungen belegt werden (Eshel & Klein, 1981; Beneson & Dweck, 1986; Weinert & Stefanek, 1997; Wigfield et al., 1997). Helmke (1991) präsentierte Befunde zum Entwicklungsverlauf der Höhe

und der Genauigkeit der Selbsteinschätzung aus einer deutschen Längsschnittuntersuchung mit sehr grosser Stichprobe, die ab 1988 über Jahre im Grossraum München durchgeführt wurde. Es sind zwei sich überlappende Untersuchungsreihen. Die uns interessierenden Befunde, welche vom Kindergarten bis zur zweiten Klasse erhoben wurden, umfassen aber nur eine Stichprobe von 118 Kindern (Tab.:1-2-1). Die Befunde beziehen sich auf den Fachbereich Mathematik.

Tabelle 1-2-1: Mittelwerte der Einschätzhöhe
im Fach Mathematik (Helmke, 1991)

Klasse	KG	1.Kl.	2.Kl.	3.Kl.
M	2.2	3.1	1.9	2.3

(-5: am Schlechtesten; 5: am Besten)

In allen Jahren überschätzen sich die Kinder im Durchschnitt (positive Abweichung vom Mittelwert 0). Es findet sich kein linearer Abfall der durchschnittlichen Einschätzhöhe: Vom Kindergarten zur ersten Klasse, sowie von der zweiten zur dritten Klasse ist die durchschnittliche Einschätzhöhe im Fachbereich Mathematik angestiegen. Der Mittelwert im ersten Schuljahr unterscheidet sich hochsignifikant von demjenigen im Kindergarten und der zweiten und dritten Klasse. Der Befund, dass die Einschätzhöhe während der Schuleingangsphase nicht kontinuierlich abnimmt, wird von einer älteren, querschnittlichen Untersuchung (Beneson & Dweck, 1986) bestätigt, die ebenfalls Kindergarten und die beiden ersten Schuljahre untersuchte. Allerdings sind pro Jahr nur 27 - 37 Kinder untersucht worden und die Befunde beziehen sich auf die Frage nach der allgemeinen Schulleistungsfähigkeit (Tab.1-2-2).

Tabelle 1-2-2: Mittelwerte der Einschätzung allgemeiner
Schulleistungsfähigkeit (Beneson & Dweck, 1986)

Klasse	KG	1.Kl.	2.Kl.	4.Kl.
Mw.	4.7	4.9	4.4	4.0

(1= am Schlechtesten; 5 =am Besten)

Die durchschnittlichen Mittelwerte sind sehr hoch, in allen Jahren überschätzen (Werte höher als 2.5) sich die Kinder signifikant. Der leichte Anstieg der Einschätzung vom Kindergarten zur ersten Klasse ist, im Gegensatz zu Helmkes Befunden, hier nicht signifikant.

1.3. Probleme mit dem diskontinuierlichen Verlauf

Gemeinsam ist diesen beiden Befunden die Diskontinuität des Verlaufs der Einschätzungshöhe, der Effekt ist bei Beneson und Dweck allerdings nicht signifikant. Wie erklärt sich der signifikante Befund von Helmke? - Er vermutet, dass diese Abweichung mit dem pädagogischen Schonraum, den das erste Schuljahr bietet, zusammenhängen könnte. Zudem könne es sein, dass in der ersten Klasse viele Kinder sich selbst für die „Besten“ hielten, weil sie die Leistungen der Mitschülerinnen noch gar nicht so gut kennen (Messpunkt im erstem Schuljahr ca. 8 Wochen nach Schulbeginn). Es kommt dazu, dass die 118 Kinder einen drei Jahre dauernden Kindergarten besucht hatten. Die Erhebung der Daten im Kindergarten umfasste im Durchschnitt nur einen Drittel aller Kindergartenkinder einer Klasse, die anderen Kinder waren ein- und zwei Jahre jünger. Damit ist die Jahrgangsvergleichsklasse („Wie gut bist du, im Vergleich zu den anderen?“) relativ klein und erst noch labil, denn wenn sich das Kind nun am Rangplatz einschätzen muss, ist diese Vergleichsgruppe viel kleiner als ein Jahr später, denn in der ersten Klasse definiert es seinen Rangplatz aufgrund von mehr als 20 Kindern des gleichen Jahrgangs. Ungünstig ist das, weil zum Messzeitpunkt in der ersten Klasse die soziale Bezugsnorm schon für einige Kinder bedeutsam sein kann: Wechselt die Bezugsgruppe stark, kann eine Einschätzung für das Kind schwieriger sein als in späteren Jahren, wenn die Bezugsgruppe über längere Zeit relativ stabil bleibt. Für einen jahresweisen Vergleich der Genauigkeit der Einschätzung der eigenen Rangposition ist eine Konstanthaltung der Bezugsgruppe wichtig.

Den leichten Anstieg von der zweiten zur dritten Klasse (nicht signifikant) erklärt sich der Autor mit dem erneuten Wechsel der Lehrkraft zu Beginn des dritten Schuljahres: Die neue Lehrperson tendiere zu einer über - optimistischen Rückmeldekultur in der ersten Hälfte des Schuljahres. Das kann sich natürlich auswirken, wenn die Erhebung relativ kurz nach Schuljahresbeginn erfolgt, wie im vorliegenden Fall. Besser wäre ein Messzeitpunkt, der nach einigen verbindlichen Leistungsrückmeldungen erfolgt und zum möglichst gleichen Zeitpunkt gegen Ende des Schuljahres stattfindet.

Gemäss den Befunden von Tiedemann und Faber ist die Wirkung von Rückmeldungen durch die Lehrkraft für die Selbsteinschätzung im Schuleingangsalter allerdings eher gering (Tiedemann & Faber, 1995; 1994). Es muss also ein anderer Grund hinter dem Anstieg der Einschätzungshöhe vom Kindergarten zur ersten Klasse vermutet werden. Die grosse Kompetenzzunahme in Mathematik im ersten Schuljahr könnte dahinter stehen, weil die intraindividuelle Bezugsnorm ja die wichtigste Bezugsnorm ist. Es würde sich deshalb lohnen neben der Mathematik einen Lernbereich zu untersuchen, bei dem der Kompetenzzuwachs vom Kindergarten zur zweiten Klasse gleichmässiger verläuft. Geeignet dafür ist der Sport, denn

im Alter zwischen 5 und 9 Jahren sind die Leistungszuwächse im Sport in vielen Fertigungsbereichen gleichbleibend. Die vorliegende Untersuchung vergleicht deshalb die Befunde im Leistungsbereich der Mathematik, die genau bei Schulbeginn das Gefühl eines grossen Leistungszuwachses zulässt, mit Einschätzungen im Leistungsbereich Sport. Dort werden die Kinder schon im Kindergarten systematisch geschult und die Lernziele stammen aus dem gleichen Lehrmittel wie in der Schule. Die Erhebungen wurden mit einer maximalen Schwankungsbreite von 2 Wochen jeweils im Mai /Juni durchgeführt. Dieser Zeitpunkt ist günstig, weil die Kinder zu diesem Zeitpunkt rund 32 von 38 Schulwochen zusammen verbracht haben.

1.4 Die Fragestellungen dieser Arbeit

Ich möchte die Frage beantworten, wie sich die Selbsteinschätzung in den Leistungsbereichen Mathematik und Sport vom Kindergarten bis zum Ende der zweiten Klasse entwickelt. Dabei geht es zunächst um die Höhe der Selbsteinschätzung, verglichen mit den tatsächlichen Leistungen. Kinder können sich unter- oder über- oder auch richtig einschätzen. Es ist interessiert darum die Frage, ob sie sich im Durchschnitt unter-, über- oder richtig einschätzen. Daran schliesst sich die Frage an, wer sich allenfalls unter- oder überschätzt, denn möglicherweise gibt es Geschlechts-, Milieu- und Altersunterschiede. Zusätzlich zur Frage der Einschätzhöhe soll die Frage der klassenrelativen Genauigkeit untersucht werden. Damit ist die Korrelation zwischen Selbsteinschätzung und tatsächlicher Leistung gemeint. Es ist ja denkbar, dass sich alle Kinder einer Klasse um einen bestimmten Betrag überschätzen, in ihren Schätzungen aber doch die Rangreihenfolge der tatsächlichen Leistungen abbilden. Wahrscheinlich ist, dass diese Korrelationen relativ gering sind und allenfalls mit zunehmender Beschulung zunehmen. Mit der Auswahl von Mathematik und Sport ergeben sich gute Vergleichsmöglichkeiten mit anderen Studien. Die Wahl von zwei Fachbereichen, bei denen häufig vermutet wird, dass Knaben bessere Leistungen erbringen, ist auch aus der Genderperspektive interessant. Knaben und Mädchen erbringen in diesem Alter nämlich in beiden Bereichen gleiche Leistungen.

Die Fragestellungen sind entwicklungspsychologisch und schulpädagogisch motiviert. Die Entwicklungspsychologie legt nahe, dass die Selbsteinschätzung im Laufe dieser Entwicklungsspanne realistischer wird. Es ist aber auch eine Tatsache, dass gerade die jüngeren Schulkinder grosse Lernfortschritte machen. Sofern sie diese wahrnehmen, sollte das Niveau der Selbsteinschätzung über die Jahre überdurchschnittlich hoch bleiben, aber doch vielleicht realistisch werden, weil sich die Kinder in der Schule dauernd mit anderen ver-

gleichen und weil die Lehrkräfte mittelfristig eine realistische Selbsteinschätzung auch pädagogisch fördern. Die Schätzgenauigkeit hingegen, das heisst die Korrelation der Leistung mit der Einschätzung, dürfte über diese Altersspanne zunehmen.

1.5 Zum Inhalt und Aufbau des Theorieteils

Das Kapitel 2.1 ist interdisziplinär angelegt und beginnt mit sozialwissenschaftlichen Überlegungen zur wachsenden Bedeutung der Fähigkeit, sich selbst einzuschätzen. Das hat auf die Schule eingewirkt, beispielsweise auf die Entwicklung von Lehrplänen, Unterrichtsdidaktik und Lehrmitteln. Dann geht es um die weiteren Einflussfaktoren, die auf die Höhe der Selbsteinschätzung und deren Verlauf Einfluss nehmen können. Diskutiert wird dann, wie offene Lernsituationen die Förderung der Selbsteinschätzungsfähigkeit zum Ziel haben und wie sie sich auf leistungsschwächere Kinder auswirken können.

Eine angemessene Selbsteinschätzung ist erst möglich, wenn die entsprechenden Kognitionen entwickelt sind: Wo steht das Kind bei Beginn der Schuleingangsphase, wo an deren Ende? Das ist das Thema des nächsten Theoriekapitels. Die Selbstkonzeptforschung hat Modelle entwickelt, die zeigen, in welcher Abfolge die nötigen kognitiven Strukturen beim Kind aufgebaut werden. Allgemein geht man heute davon aus, dass dieser Entwicklungsprozess mit etwa 10 bis 12 Jahren abgeschlossen ist (Flammer, 1996; Oerter, 1995).

Anschliessend folgt die Diskussion von Befunden, die sich mit der Entwicklung des Fähigkeitsselbstbildes im Schuleingangsalter befassen. Das führt uns zur Klärung der Frage, wo das Kind bezüglich seiner kognitiven Entwicklung im Kindergartenalter steht: Es weiss in einigen Leistungsbereichen um sein Können oder den eigenen Misserfolg. Es kann zudem Erfolgskriterien folgerichtig anwenden (Flammer, 1996).

Während der Schuleingangsphase erlernt das Kind, dass Fähigkeit und Anstrengung Unterschiedliches meinen und zusammen die Leistung erhöhen können: Der Aufbau und die Ausdifferenzierung des Attributionskonzeptes wird in einem weiteren Kapitel dargestellt, um zu verstehen, welche kognitiven Voraussetzungen gegeben sein müssen, damit Kinder die verschiedenen Leistungsrückmeldungen richtig verstehen und interpretieren können.

Mangelndes Selbstvertrauen wird in der Gender-Forschung häufig als Grund angeführt, warum Frauen im Beruf seltener aufsteigen und Kaderstellen besetzen. Ein Kapitel bringt Befunde zur Entwicklung des Selbstkonzeptes im Grundschulalter bei Mädchen und Knaben. Haben Mädchen schon in der Schuleingangsphase die Tendenz, sich leistungsmässig weniger zuzutrauen als Knaben?

Ein letztes Interesse gilt der Frage, ob sich eher ländlich aufgewachsene Kinder leistungsmässig weniger hoch einschätzen, als solche aus urbanen Wohngebieten. Bei der Auswahl der Klassen wurde deshalb darauf geachtet, genügend Klassen aus dem ländlichen Raum für die Untersuchung zu gewinnen.

2. Die Entwicklung der Fähigkeitsselbsteinschätzung im Schuleingangsalter

2.1 Einleitung

Es sind viele Variablen, die die Entwicklung der Selbsteinschätzung beeinflussen können. Im Kap.2.2 werden sie diskutiert. Es beginnt mit den gesellschaftlichen Hintergründen, die die Entwicklung und die Wahrnehmung des Selbst beeinflussen. Was die Gesellschaft fordert, wird von der Schule früher oder später gefördert. Das ist auch in diesem Fall so. Es wird deshalb hergeleitet, wie die Schule bis in den Unterricht hinein das kindliche Bewusstsein um die eigene Leistungswahrnehmung fördert.

Lernen findet in der Schule in einem sozialen Rahmen statt, auch die Mitlernenden können das individuelle Leistungsselbstbild beeinflussen. Im Schuleingangsalter ist das allerdings nur begrenzt möglich, wie wir sehen werden.

Zwischen dem sechsten und zehnten Lebensjahr macht die kognitive Entwicklung grosse Fortschritte. Das sich entwickelnde Denken hat einen direkten Einfluss auf die Möglichkeiten, sich selber realistisch wahrzunehmen und zu beurteilen. Von allen Variablen ist das für die Selbstbildentwicklung im Schuleingangsalter die bedeutendste, entsprechend differenziert wird dieser Aspekt ausgeleuchtet.

Die Darstellung der wichtigsten Befunde aus der Forschungstradition zur Entwicklung der Einschätzhöhe in schulischen Leistungsbereichen führt uns (in Kap.2.4) bis zu erst kürzlich an der Universität Bern erhobenen Daten, die auf den Instrumenten beruhen, wie sie in der vorliegenden Arbeit benützt werden. Weiter geht es mit Kapitel 2.5, der Darstellung von Befunden zur unterschiedlichen Selbsteinschätzung von Jungen und Mädchen. Dann folgen die Hypothesen, womit der Theorieteil abgeschlossen ist.

2.2 Determinanten der Selbsteinschätzung

2.2.1 Die Determinanten der Selbsteinschätzung im Überblick

Im ersten Kapitel wurde dargelegt, dass die Befundlage zum Verlauf und der Genauigkeit der Höhe der Selbsteinschätzung diskontinuierlich sein kann und eine Abnahme während der Schuleingangsphase nur bedingt belegt ist: Einen wichtigen Einfluss auf den Verlauf hat der Übergang des Kindes von einer Vorschulinstitution (Kindergarten) zur Volksschule: Dadurch werden häufig Klassen neu gebildet und neue Leistungsbereiche kommen dazu. Das erschwert den Kindern, die eigene Leistungsfähigkeit einzuschätzen. Genaueres über den tatsächlichen Verlauf der Selbsteinschätzung kann erfahren werden, wenn es gelingt, diese Faktoren zu kontrollieren, respektive deren mögliche Wirkung tief zu halten. Trotzdem: Die Befunde von Helmke bestätigen, dass sich die Kinder aller Leistungsstärken im Kindergarten und während den ersten Schuljahren massiv überschätzen. Sehr gut belegt ist auch, dass gegen Ende der Grundschulzeit, also im sechsten Schuljahr, die klassendurchschnittliche Überschätzung verglichen mit dem ersten Schuljahr deutlich abgenommen hat (Helmke, 1992). Sicher überschätzen sich Kinder im Schuleingangsalter, noch unsicher ist, ob diese Überschätzung in diesen Jahren stabil bleibt oder zu -, respektive abnimmt. Die Hauptgründe für die Überschätzung der eigenen Leistungsfähigkeit und die sehr tiefen Korrelationen während der Schuleingangsphase werden in noch nicht ausreichend ausdifferenzierten kognitiven Kompetenzen vermutet (Nicholls, 1978; Parsons & Ruble, 1977; Shaklee, 1976). Insbesondere der Aufbau des Ursachenzuschreibungsmodelles und dessen Verständnis sei noch nicht ausreichend (Heckhausen 1980; Nicholls, 1978; Nicholls & Miller, 1984; Stipek & Tannatt, 1984).

Diese Annahmen können allerdings die im einleitenden Kapitel gezeigten diskontinuierlichen Verläufe der Einschätzhöhe nicht erklären. Wichtige Einflussvariablen, Determinanten des Fähigkeitsselbstbildes also, variieren in der Grundschulzeit: So führt beispielsweise der Wechsel der Lehrperson regelmässig zu einem Anstieg (statt einem Absinken) der Überschätzung (Helmke, 1999). Im folgenden werden zuerst mögliche wichtige Determinanten des Fähigkeitsselbstbildes zusammenfassend dargestellt, die einen Einfluss auf die Einschätzhöhe haben können. Es handelt sich dabei um Determinanten, die während der Schuleingangsphase von Bedeutung sind, Tab.2-2-1 fasst sie zusammen, sie sind angelehnt an die Überlegungen von Helmke (1992).

Tabelle 2-2-1: Determinanten des Fähigkeitsselfbildes (angelehnt an Helmke, 1992)

<p>Determinanten des gesellschaftlich/schulischen Kontextes</p> <p>1.) Der gesellschaftliche Kontext</p> <p>2.) Der normative Kontext der Schule</p> <p>3.) Selbstgesteuerte Unterrichtsformen</p> <p>4.) Beurteilung und Schwierigkeitsgrad der Lernaufgaben</p> <p>5.) Die Art und der Schwierigkeitsgrad der Lernaufgaben</p>
<p>Determinanten der Schulklasse</p> <p>6.) Die Leistungsstärke der Mitschülerinnen und –schüler</p> <p>7.) Die Bedeutung der Leistungsnorm für das Prestige</p>
<p>Determinanten des Individuums</p> <p>8.) Bisherige Leistung und persönlicher Standard</p> <p>9.) Der Wandel in der Bezugsnormorientierung</p> <p>10.) Die Lernfreude des Kindes</p> <p>11.) Entwicklung und Ausdifferenzierung des Attr.konzepts</p>

2.2.2 Determinanten des gesellschaftlich / schulischen Kontextes

2.2.2.1 Der gesellschaftliche Kontext:

Die steigende Bedeutung der Selbstverwirklichung

Zu Beginn des einundzwanzigsten Jahrhunderts gehen wir im euro - amerikanischen Raum davon aus, dass Menschen eigenverantwortlich ihren Lebenslauf bestimmen können und deren Wohlbefinden vom Ausmass dieser Selbstverwirklichung beeinflusst ist (Grob, 1996; Flammer, 1990). Anstelle von gesellschaftlich vorgegebenen, für alle gültigen Lebenszielen, werden Lebensaufgaben wichtiger, die sich das Individuum selbst stellt. Das Individuum wird damit zum „Schmied seines eigenen Glücks“. Dieser gesellschaftlich-kulturellen Entwicklung ist das Individuum ausgesetzt, unabhängig davon, ob es das möchte oder nicht. Sich selbst zu verwirklichen ist eine gesellschaftliche Norm geworden.

Diese Entwicklung bedingt, dass Menschen Kontrolle über den Verlauf ihres Lebens gewinnen wollen, um ihrer individuellen Entwicklung den gewünschten Weg zu bahnen (Ban-

dura, 1995). Allerdings gibt es auch Grenzen für dieses Kontrollbedürfnis, denn das Leben lehrt, dass man gar nicht alles kontrollieren kann (Flammer, 1990). Kontrolle kann ausgeübt werden, wenn man selbst über Fähigkeiten oder Kompetenzen verfügt um die Umwelt so zu beeinflussen, dass sie den eigenen Zielen förderlich ist, oder, indem man diese von anderen Personen „ausleiht“ (was beispielsweise kleine Kinder tun, wenn sie die grössere Schwester zum eigenen Vorteil mobilisieren). Allerdings möchten schon Kinder von fremder Hilfe unabhängig werden und auf eigene Kompetenzen und Fähigkeiten aufbauen um Kontrolle zu erreichen (Flammer, 1996). Das Bedürfnis zum Kontrollgewinn ist schon in der Kindheit vorhanden.

Kulturübergreifende Studien zeigen, dass die eingangs erwähnte Abhängigkeit des Wohlbefindens vom Grad der Selbstverwirklichung nicht in allen Kulturen so ausgeprägt ist, wie im euro-amerikanischen Kulturraum: In einer Vergleichsstudie zwischen Personen aus den USA und Japan fand sich bei den Personen aus Japan ein deutlich geringerer Zusammenhang zwischen Wohlbefinden und dem Grad der Selbstverwirklichung (Markus, Kitayama, Matsumoto & Norasakunit, 1997). Flammer et al. (1995) fanden in einer kulturvergleichenden Studie zwischen Jugendlichen aus der Schweiz und Japan ebenfalls signifikante Unterschiede im Umgang mit Kontrollverlust. Die japanischen Jugendlichen waren viel eher bereit, sich mit dem Kontrollverlust abzufinden. Grundsätzlich wollten sie aber Kontrollverluste ebenso stark vermeiden, wie die Jugendlichen in der Schweiz. Während das Kontrollbedürfnis in Bezug auf das eigene Leben in allen Menschen vorhanden ist (Flammer, 1990), scheint dessen Ausprägung und Intensität kulturell gebunden zu sein. Anders gesagt, bei der hohen Bedeutung der Selbstverwirklichung für unser Wohlbefinden handelt es sich um ein Ergebnis euro-amerikanischer Entwicklung, die zur kulturellen Norm geworden ist. - Werte und Normen an die nachfolgende Generation weiterzuvermitteln ist eine Kernaufgabe der Schule (Reble, 1981; Fend, 1974). Die wachsende Bedeutung der Selbstzentrierung und die daraus folgende Wichtigkeit des Selbstkonzeptes ist eine kulturelle Entwicklung, die schon seit längerem Wirkungen auf die Erziehungsziele der Schule ausübt (Giesecke, 1988), sie wird im folgenden nachgezeichnet.

2.2.2.2 Der normative Kontext: Die Förderung eines positiven Selbstkonzeptes in der Schule dargestellt am Beispiel des Kantons Bern

Im Schweizer Kanton Bern finden sich in den Lehrplänen bis in die siebziger Jahre (Erziehungsdirektion des Kantons Bern, 1973) kaum Hinweise auf Bildungsziele, die über die Sachkompetenzförderung hinausgehen. Eine Ausnahme bildete allenfalls die explizit formu-

lierte Erwartung, dass Kinder in die christliche Tradition protestantischer Prägung eingeführt werden sollen. Mit dem 1983 eingeführten Lehrplan (Erziehungsdirektion des Kantons Bern, 1982) kamen neu auch explizit formulierte Ziele zur Förderung der Sozial- und Selbstkompetenz hinzu, allerdings nur im Sinne allgemeiner Leitideen. Der aktuelle bernische Lehrplan (Erziehungsdirektion des Kantons Bern, 1995) geht diesbezüglich tiefer und differenziert die Umsetzung der Förderung der Sozial- und Selbstkompetenz durch Aussagen in den Leitideen und durch entsprechende Zielsetzungen in geeigneten Fachbereichen, beispielsweise im Fach Mathematik oder in „Natur-Mensch-Mitwelt“. In den Leitideen zur Selbstkompetenz (Erziehungsdirektion des Kantons Bern, 1995, Leitideen 2) finden sich folgende Zitate:

„Selbständigkeit äussert sich in der Fähigkeit, sich ein eigenes Urteil zu bilden.... . Selbständigkeit setzt Selbstvertrauen voraus,....“

„Zur Persönlichkeitsentwicklung gehört auch ein bestimmtes Mass an Selbstkritik.“

„Sie fördert die Fähigkeit der Schülerinnen und Schüler, sich selber zu beurteilen. So lernen diese, Verantwortung für das eigene Lernen und die eigene Schullaufbahn zu übernehmen.“

In den Allgemeinen Bestimmungen des Lehrplans findet sich folgende Aussage (Erziehungsdirektion des Kantons Bern, 1995, AHB 20):

„Schülerinnen und Schülern sollen lernen, ihr eigenes Arbeiten zu beurteilen. Die Fremdbeurteilung kann in zunehmendem Masse durch die Selbstbeurteilung ersetzt werden.“

Bis in die Bildungsziele der Schule lässt sich die zunehmende Bedeutung der Selbstzentrierung also nachzeichnen. Das Kind soll durch die Reflexion eigenen Lernens, durch die Beurteilung eigenen Misserfolgs und Erfolgs selbstkritisch werden. Selbstkritikfähigkeit ist in den Augen des Lehrplans ein wesentliches Merkmal der Persönlichkeitsentwicklung. Was am Beispiel der Lehrplanentwicklung im Kanton Bern aufgezeigt wurde, gilt zunehmend für alle (deutsch-) schweizerischen Lehrpläne: Die Kantone Thurgau, St.Gallen, Solothurn, Luzern und Schaffhausen benennen die Selbstkompetenzförderung als zentrales Ziel der Schule. Zürich, Appenzell Innerrhoden erwähnen sie sehr summarisch. Allgemeine Bildungs- und Erziehungsziele verändern sich auch in Deutschlands Schulen in die gleiche Richtung (Faust-Siehl et al., 1996).

Was für Lehrpläne gilt, kann auch in der Unterrichtsdidaktik beobachtet werden: Verbreitete, in der letzten Dekade in der Schweiz herausgegebene Standardwerke zur Unterrichtsmethodik für die Volksschule betonen die zunehmend wichtiger werdenden Lernfor-

men, die das individuell und selbstgesteuert lernende Kind ins Zentrum rücken schon seit längerem (Gasser, 1995; Ruf & Gallin, 1990). Dieser Tendenz folgend haben in der Volksschule Unterrichtsformen Einzug gehalten, die eine hohe Selbststeuerung des Lernprozesses durch die Kinder unterstützen und herausfordern. Es sind dies Unterrichtskonzepte, welche im deutschsprachigen Raum unter den Namen „Wochenplan-“, „Projekt-“, und „Werkstattunterricht“ bekannt geworden sind und auf die Reformpädagogik am Beginn des zwanzigsten Jahrhunderts zurückgehen (Catani, 1999). Zentrales und gemeinsames Ziel dieser Unterrichtsformen ist es, Kinder selbständig und auf eigenen Wegen aus einer Vielzahl von Lernaufgaben einen wirksamen, optimalen Lernweg herausarbeiten zu lassen und diesen Lernweg zu gehen. In den Didaktiklehrmitteln wird abgeleitet, dass Kinder dadurch motivierter sind und deshalb mehr lernen.

2.2.2.3 Selbstgesteuerte Unterrichtsformen setzen Selbstvertrauen voraus

Ein Problem aller dieser Unterrichtskonzepte ist, dass sie das autonome, neugierige und handlungswillige Kind bereits voraussetzen, also davon ausgehen, Kinder würden auf offene Lernsituationen mit höherer Motivation und Leistung reagieren. Dabei bleiben die Befunde aus der Psychologie weitgehend ausgeblendet, die zeigen, dass Lernsituationen in der Volksschule auch die gegenteilige Wirkung haben können, insbesondere wenn Kinder zu Strategien des Selbstwertschutzes greifen (Meyer, 1984; Nicholls, 1984): Diese Strategien dienen dazu, eine Abwertung des Selbstwertes zu vermeiden, etwa wenn ein Kind annimmt, dass es eine bestimmte, wichtige Leistung nicht erbringen kann (hoher Anreiz, niedrige Handlungs-Ergebnis-Erwartung) und gleichzeitig befürchtet, dass dieses Versagen peinlich ist und zu einem Imageverlust führt. In einem Leistungswettbewerb, wie er in der Schule stattfindet ist Misserfolg, beispielsweise in Form schlechter Noten oder ungünstiger sozialer Vergleichsprozesse („ich bin ja der langsamste“) unvermeidlich. Kinder, die häufig solche Situationen erleben, stellen ihre Anstrengung bezüglich der Erreichung des Lernzieles mehr und mehr ein und konzentrieren sich darauf, Misserfolg zu vermeiden (Jagacinski & Nicholls, 1980). Diese Kinder können zur Verweigerung einer bestimmte Leistung neigen („Für die Sekundarschulprüfung melde ich mich gar nicht an“), sie strengen sich also nicht an (Heckhausen & Rheinberg, 1980). Andere werten die Lernaufgabe ab („Das kann man im Leben gar nie brauchen“), betonen also eine mangelnde Lernmotivation, usw. Andere Kinder neigen dazu, sich eine viel zu leichte Aufgabe vorzunehmen, damit gar kein Risiko auf Misserfolg vorhanden ist (aber vermutlich auch kaum etwas zu lernen ist) oder sie bearbeiten eine Lernaufgabe, die eigentlich fast niemand kann. Dann kann es auch vorkommen, dass Kinder

einen Misserfolg in seiner Bedeutung abwerten („Französisch brauche ich ja später sowieso nicht“). Das Problem all dieser defensiven Strategien ist, werden sie im Übermass eingesetzt, führen sie tatsächlich zum befürchteten Misserfolg (Helmke, 1992): Mangelndes Selbstvertrauen kann das selbstgesteuerte Lernen in ungünstige Bahnen lenken. Auf diese Gefahren machen Standardwerke der Pädagogischen Psychologie aufmerksam (Mietzel, 1993; Weidenmann & Krapp, 1986).

Wie kommt es, dass die didaktische Literatur für die Volksschule dieses Wissen zur „Psychologie des Lerner“ nicht aufgenommen hat? Vermutlich aus zwei Gründen: Erstens begründet die Volksschuldidaktik ihre unterrichtsmethodischen Konzepte im Wesentlichen mit ihrer Bewährung in praktischen Versuchen. Sie verweist dabei auf einzelne pädagogische Leitfiguren, die damit einen hohen Lernerfolg erzielten (beispielsweise Freinet in den dreissiger Jahren mit seinem Werkstattunterricht). Dabei werden die Rahmenbedingungen der meist schon länger zurückliegenden Versuche nicht diskutiert und die Versuche wurden in den meisten Fällen kaum wissenschaftlich begleitet. Nachteile von Unterrichtsformen bleiben in der Literatur zur Didaktik praktisch ausgeblendet (Catani, 1999). Daraus entsteht dann im Volksschulalltag häufig eine zu hohe Erwartungshaltung gegenüber neueren Unterrichtsformen. Zweitens spielt die Psychologie bezüglich der theoretischen Fundierung von Unterrichtsformen nur eine Nebenrolle innerhalb der Grundausbildung von Lehrpersonen: Sie wird höchstens dann zitiert, wenn sie die Didaktik bestätigt (Meier, 1999).

Zentrales Ziel dieser heute stark verbreiteten offenen Unterrichtsformen ist es, Kinder selbständig und auf eigenen Wegen aus einer Vielzahl von Lernaufgaben einen wirksamen, optimalen Lernweg herausarbeiten zu lassen. Sie sollen ihr Lernen reflektieren und ihre Lernergebnisse selbst korrigieren. Die Kinder arbeiten im Schulzimmer zwar selbständig und selbstgesteuert, jedoch in stetem Kontakt zu ihren Mitschülerinnen und -schülern: Sie sehen schon nach kurzer Zeit, wenn andere Kinder mehr Lernaufgaben zu bewältigen vermögen oder Arbeiten planen, die als schwierig eingestuft werden müssen. Sie stehen also mit ihrer Tätigkeit in dauerndem sozialen Vergleich. Gezielte Beobachtungen¹ in offenen Unterrichtssituationen bestätigten die von Meyer (1984) und Nicholls (1984) oben berichteten Befunde, die diese in geführtem, lehrerzentrierten Unterricht beobachteten: Die einen Kinder gehen schnell und speditiv an die Arbeit, andere Kinder trödeln, schauen zuerst, was andere (vorwiegend leistungsstarke Kinder) tun, informieren sich bei ihren Nachbarn oder wollen jede Entscheidung durch Rückfragen bei den Lehrkräften absichern.

¹ 120 Kinder wurden systematisch in offenen Lernsituationen bezüglich den von Meyer (1984) genannten Verhaltensweisen beobachtet (im Rahmen der Lehrergrundausbildung in Spiez, Ergebnisse nicht publiziert).

Werkstatt- und Wochenplanunterricht (die hauptsächlichsten Formen offenen Unterrichts) sind von der methodischen Anlage her so konzipiert, dass alle Merkmale ausgeprägter Leistungssituationen zum Tragen kommen, wie sie von Heckhausen (1974) definiert wurden:

- a) die Handlungen des Kindes müssen zu einem überprüfbaren Resultat führen
- b) die richtige Lösung der Aufgaben ist nicht mit Sicherheit gegeben
- c) es braucht einen Gütemassstab, respektive eine Bezugsnorm
- d) der Handelnde muss sich für das Resultat selbst verantwortlich fühlen

Im Werkstattunterricht sind diese Merkmale alle gegeben: die Kinder bearbeiten einen schriftlichen Auftrag alleine, führen anschliessend die Korrektur selber durch (die Lehrkraft macht Stichproben). Aus einem Angebot von vielen Lernaufgaben können sie auswählen, wobei der Schwierigkeitsgrad der einzelnen Aufgabe durch die Lehrkraft deutlich angegeben wird. Zudem gibt es meistens eine Zeitvorgabe (beispielsweise sechs Lektionen Zeit) und einen Gütemassstab, welcher das Minimalziel definiert: Also beispielsweise 8 von 20 Lernaufgaben lösen. Den Arbeitsstand der einzelnen Kinder dokumentiert eine Klassentabelle, auf der alle Kinder ihre bearbeiteten Aufträge eintragen, sodass für Lehrkraft und Mitschüler leicht ersichtlich ist, wer wieviel bearbeitet hat. Die Kinder können also nach einigen Lektionen durchaus realisieren, dass sie eher schnell oder langsam sind oder dass signifikante Andere (beispielsweise der Klassenprimus) auch schwierige Aufgaben ohne Hilfe bewältigen können (soziale Vergleichsnorm). Da die Kinder bei auftauchenden Schwierigkeiten die Lehrkraft konsultieren können, bleiben sie bei einem allfälligen Scheitern bei der Bewältigung des Auftrages alleine verantwortlich. Unterrichtsformen, wie sie der Werkstattunterricht darstellt, werden in der Volksschule immer häufiger und die vorerwähnten Lehrplanziele können ohne diese und ähnliche Unterrichtsformen gar nicht mehr erreicht werden. Auch neue Lehrmittel in den zentralen Schulfächern Deutsch („Sprachfenster“, Staatlicher Lehrmittelverlag, 2000), Mathematik („Das Zahlenbuch“, Klett-Verlag, 1995) und Natur-Mensch-Mitwelt, welche fast 70 % der gesamten Unterrichtszeit beanspruchen, bauen zentral auf dem selbstgesteuerten Lernen auf und führen damit zu Leistungssituationen mit den von Heckhausen beschriebenen Merkmalen.

Selbststeuerung bedeutet häufig, sich zwischen verschiedenen Lernalternativen zu entscheiden. Am Beispiel des Werkstattunterrichts zeigt sich, wie das Leistungselbstbild diesen Entscheidungsprozess und letztlich damit den Lernerfolg des Kindes mitsteuert. Eine sogenannte „Werkstatt“ besteht aus einer Anzahl von einzelnen Arbeitsposten, die ein Kind möglichst selbständig auswählen, bearbeiten und korrigieren soll. Die „Werkstatt“ trägt einen

Titel, beispielsweise „Tiere im Jahreslauf“ als Umsetzung des Lehrplanziels „Naturerscheinungen – Naturbegegnung“ (Lehrplan für den Kanton Bern, 1995, NMM 16). Die „Werkstatt“ besteht aus möglicherweise 15 Einzelaufgaben, die je ein bestimmtes Lernziel abschliessend erarbeiten. Um eine solche Einzelaufgabe zu bearbeiten, braucht das Kind 20 – 60 Minuten. Das Deckblatt zu jeder Einzelaufgabe nennt die vermutete Arbeitszeit, den Schwierigkeitsgrad der Einzelaufgabe, das Lernziel und das Vorgehen: Im Grundsatz geschieht damit das Gleiche, wie wenn die Lehrkraft eine Einzelaufgabe mit allen Kindern zusammen erarbeiten würde. Die pädagogische Hoffnung besteht darin, dass das Kind durch die Selbststeuerung des Lernprozesses selbständiger wird: Vor Allem soll das Kind lernen, was es sich zumuten darf: Es wählt ja selber, ob es leichte, mittlere oder schwierige Aufgaben lösen will. Ist eine Einzelaufgabe bewältigt, trägt es sich auf einer Klassenübersicht ein: dort ergibt sich die Möglichkeit zu sehen, wie viel und welche schwierigen Aufgaben andere der Klasse schon geleistet haben. Das Lernen findet im Werkstattunterricht unter typischen Wettbewerbsbedingungen statt: Alle kennen Ziele, Arbeitsbedingungen und Schwierigkeitsgrad der Lernherausforderungen und können direkt sehen, in welchem Mass sie selbst und andere diesen Herausforderungen gewachsen sind. Steht das Kind nun vor der Situation, eine neue Einzelaufgabe auszuwählen, kommen kognitive Prozesse in Gang, bei denen das Vertrauen in die eigene Leistungsfähigkeit eine entscheidende Rolle spielt. An einem Beispiel wird nun der Verlauf einer solchen Situation nachgezeichnet (nach Heckhausen, 1989, 469): Das Kind steht vor der Liste aller 15 Einzelaufgaben und muss entscheiden, welche Aufgabe es jetzt zur Bearbeitung auswählt.

Situation – Ergebnis – Erwartung:

Die erste Frage, die sich das lernende Kind nun stellt, ist diejenige, ob in der Einzelaufgabe eine echte neue Erkenntnis steckt, die man noch erwerben sollte. Ist das nicht der Fall, kann die Einzelaufgabe weggelassen werden (in „Werkstätten“ ist dies normal, denn etwa ein Drittel der Einzelaufgaben braucht nicht gelöst zu werden, um die Grundanforderungen zu erfüllen).

Nun findet ein Kind vielleicht vier Einzelaufgaben, die eine echte Lernherausforderung darstellen: eine schwierige, eine mittlere und zwei leichte.

Handlungs – Ergebnis – Erwartung:

Nun überlegt sich die lernende Person, bei welchen Einzelaufgaben sie mit hoher Wahrscheinlichkeit zu einem Lernerfolg kommt. Um dies zu entscheiden, hat es auf dem Deckblatt

Angaben zur Schwierigkeit der Lernaufgabe: Ein erfolgreiches Kind mit gutem Selbstvertrauen wird hier vermutlich zwischen einer mittleren und einer schwierigen Aufgabe auswählen.

Handlungs – Situations - Ergebnis-Erwartung:

Nun hat die lernende Person vielleicht zwei schwierige Posten vor sich liegen und sieht, dass die verfügbare Zeit (entsprechende Angaben finden sich auf dem Deckblatt der beiden Einzelaufgaben) für die Lösung knapp ist. Vor dem knappen Zeitbudget hat sie Respekt, weil die eigene Konzentration nicht mehr als hoch beurteilt wird: Sie wählt deshalb nun doch eine mittelschwere Einzelaufgabe, die weniger Zeit beansprucht, weil ein Lernerfolg unbedingt angestrebt wird.

Ergebnis – Folge – Erwartung:

Welche Konsequenz hat nun die Auswahl dieser Einzelaufgabe? Um diese Frage zu beantworten, kommt es darauf an, welches Kriterium der lernenden Person sehr wichtig ist, es ergeben sich drei Möglichkeiten (nach Heckhausen & Rheinberg, 1980):

- 1.) Das Erreichen einer sehr guten Note wird nur möglich, wenn eine schwierige Einzelaufgabe gewählt wird (ein gutes Zeugnis ist wichtig)
- 2.) Unbedingt wichtig ist, dass man diese Einzelaufgabe erfolgreich abschliesst, damit man auf sich stolz sein kann (man will ein Erfolgserlebnis haben)
- 3.) Unbedingt wichtig ist, dass die Lehrkraft Freude über die Auswahl der Einzelaufgabe durch die lernende Person empfindet (man möchte die Lehrperson nicht enttäuschen, sondern beeindrucken)

Eine sehr gute Schülerin wird zwischen den drei Möglichkeiten nicht auswählen müssen, denn sie wird eine schwierige Aufgabe mit einiger Sicherheit lösen können. Eine eher schwächere Schülerin muss sich zwischen den Möglichkeiten entscheiden: eine einfache Aufgabe ist für sie vermutlich lösbar, bedeutet aber auch nur eine genügende Note und wenig Freude bei der Lehrperson. Damit gerät das Kind in ein Dilemma: minimalen Lernerfolg absichern oder die Lehrperson positiv überraschen? - Erst wenn dieser vierteilige Prozess nach Heckhausen abgelaufen ist, beginnt das Kind mit der Arbeit an der Einzelaufgabe. Je positiver das Ergebnis bei jedem Teilschritt ausgefallen ist, desto motivierter und entsprechend ausdauernder wird das Kind arbeiten.

Bei all diesen Überlegungen spielt das Leistungsselbstbild nun eine entscheidende Rolle: die erste Erwartung besteht ja darin, zu beurteilen, ob die Einzelaufgabe überhaupt eine Herausforderung darstellt, oder ob sie beispielsweise sehr leicht ist, weil man Ähnliches

schon mit links gelöst hat. Ein anderes Kind dagegen kommt bei der gleichen Aufgabe zum Schluss, dass eine vergleichbare Aufgabe schon einmal nur mit grosser Anstrengung gelöst werden konnte: für dieses Kind stellt diese Einzelaufgabe eine echte Herausforderung dar, es wird sich stark anstrengen müssen, um sie zu bewältigen. In beiden Fällen sind die Kinder darauf angewiesen, bereits Lebenserfahrung mit Lernherausforderungen dieser Art gehabt zu haben.

Die Handlungs – Ergebnis – Erwartung darf demgegenüber nicht zu tief sein, das Kind muss an einen Erfolg glauben, sonst ist es nicht motiviert, die Aufgabe anzupacken. Es braucht also Selbstvertrauen.

In solchen Entscheidungssituationen ist die Abweichung der Selbst- von der Fremdeinschätzung eigener Leistungsfähigkeit bedeutsam: Ist das Kind realistisch (keine Über- und Unterschätzung), dann dürfte das Kind bei angemessener Anstrengung die richtigen Einzelaufgaben auswählen und diese meist erfolgreich bewältigen. Dies ist das pädagogisch erwünschte Szenario. Hin und wieder ein Misserfolg ist dabei durchaus erwünscht, das zeigt auf, dass das Kind ein angemessenes Risikoverhalten zeigt. Überschätzt sich das Kind, dann sind Misserfolge häufig, das Kind wählt zu schwierige Aufgaben aus. Mit der Zeit wird der gehäufte Misserfolg die Anstrengungsbereitschaft lähmen. Die Lehrkraft wird intervenieren und dem Kind Einzelaufgaben zuweisen, die eine höhere Erfolgsquote nach sich ziehen. Unterschätzt sich das Kind, dann passieren ihm selten Fehler, aber es kennt damit seine Leistungsgrenze auch nicht, die Lehrperson wird es anhalten, sich schwierigere Aufgaben zuzutrauen. Im Klassenvergleich sind die Leistungen vermutlich unterdurchschnittlich.

Das Verhalten der Lehrkraft wirkt dahingehend, dass diejenigen Kinder das erwünschte Auswahlverhalten zeigen, die über eine realistische Vorstellung ihrer Leistungsfähigkeit verfügen. Hier liegt auch ein wesentliches Ziel des Werkstattunterrichts: Kinder sollen lernen, sich selbst angemessene Herausforderungen zuzutrauen. In den mittleren Grundschuljahren, ab etwa dem vierten Schuljahr (Nicholls, 1984) ist das Kind in seiner kognitiven Entwicklung so weit, dass es in der Lage ist, sich realistisch beurteilen zu können und entsprechend angemessen schwierige Aufgaben auszuwählen. Ein günstiges leistungsbezogenes Selbstvertrauen ist dafür wichtig. Der Werkstattunterricht ist auf der Schuleingangsstufe sehr beliebt bei den Lehrkräften. Kinder dieses Alters werden sich zu einem grossen Teil eher zu schwierige Aufgaben vornehmen, selten jedoch zu leichte.

Offenes, selbstgesteuertes Lernen wird von einem Teil der Schülerinnen negativ, vom anderen Teil positiv erlebt. Ein ausgeprägtes Interesse haben Kinder dann, wenn sie ein Selbstkonzept hoher Fähigkeit haben, also davon ausgehen, dass sie im entsprechenden Fachbereich gut leistungsfähig sind. Sie verfügen über gutes Selbstvertrauen. Bei tiefem Selbstvertrauen droht hingegen die Selbststeuerung des Lernprozesses so zu verlaufen,

dass Misserfolge häufig, vielleicht sogar unausweichlich sind: Ohne ein gutes Stück Selbstvertrauen kann die Selbststeuerung zu ungünstigen Lernergebnissen führen. Kinder mit hohem Selbstvertrauen haben hohe Erwartungen an sich selbst, was die Motivation positiv beeinflusst (Heckhausen, 1974). Diese Erwartungen setzen sich zusammen aus Erfahrungen mit dem eigenen Leistungswillen, den eigenen Fähigkeiten und der Einschätzung der Aufgabenschwierigkeit. Das richtige Kombinieren aller dieser Überlegungen gelingt Kindern erst mit etwa 10 Jahren richtig. In der Schuleingangsphase, also im Alter zwischen sechs und neun Jahren, sind diese Voraussetzungen nur teilweise gegeben, Kinder dieses Alters sind generell optimistischer, ihr Selbstvertrauen ist grösser, als es von der Leistung her gerechtfertigt wäre (Flammer, 1996; Oerter & Montada, 1995). Wenn Kinder in der Schuleingangsphase diese offenen Lernsituationen häufig erleben, dann dürfte es aber bei der Selbststeuerung zu Fehleinschätzungen kommen: Die Kinder neigen wahrscheinlich dazu, zu schwierige Aufgaben auszuwählen, weil sie ihre Leistungsfähigkeit überschätzen. Die auf der Unterstufe weitverbreitete Form des Werkstattunterrichts setzt sich gerade zum Ziel, die eigene Leistungsfähigkeit genauer abschätzen zu lernen. Diese Zielsetzung dürfte in diesem Alter kaum erreicht werden. Es können sich aber für schwächere Kinder zusätzliche Nachteile einstellen: Weil sie sich bei der Auswahl des Schwierigkeitsgrades ihrer Lernaufgaben überschätzen, werden die leistungsschwächeren Kinder häufig Misserfolg erfahren. Gegenüber einem Unterricht, in dem die Lehrkraft die zu lösenden Aufgaben den Kindern leistungsgerecht zuteilt, führen selbstgesteuerte Unterrichtsformen dazu, dass die Überschätzung eigener Leistungsfähigkeit kleiner ist. Als Determinante der Höhe der Selbsteinschätzung werden also offene Unterrichtsformen vermutlich zu einer insgesamt weniger hohen Überschätzung eigener Leistungsfähigkeit führen.

2.2.2.4 Beurteilung und Schwierigkeitsgrad von Lernaufgaben

Die letzte Determinante des Fähigkeitsselbstbildes, die aus dem gesellschaftlich / schulischen Kontext heraus auf die Selbsteinschätzung einwirkt, ist die Rückmeldung der Leistungsergebnisse durch die Lehrkraft. Wie Helmke (1991) zeigte (Tab. 2-2-2), sind die Leistungsrückmeldungen in Form von Noten in der Unterstufe nicht so verteilt, wie dies dann zum Selektionszeitpunkt (in vielen Bundesländern Deutschlands Ende der vierten Klasse, in der Schweiz am Ende des 6. Schuljahres) sein wird. Weil im Kindergarten, aber auch in der ersten Klasse noch keine Noten erteilt wurden, beginnt die Erhebung in der 2. Klasse

Tabelle 2-2-2: Notenverteilung bei 1300 Schülern im zweiten Schuljahr und ein Jahr später (Helmke, 1991)

	Schwache Leistung = 5	Ungenügende Leistung = 4	Genügende Leistung = 3	gute Leistung = 2	sehr gute Leistung = 1
2.Kl.	3 %	8 %	23%	38 %	25 %
3.Kl.	6%	4%	24%	45%	20%

Rund 10% der Kinder erhalten ungenügende oder schwache Noten als Leistungsrückmeldungen am Ende des Schuljahres und zwar in beiden Jahren. Stabil ist der Anteil „genügender“ Schülerinnen über beide Schuljahre, während derjenige der mit „sehr gut“ benoteten Schülerinnen und Schüler abnimmt. Die Leistungsbeurteilungen der Lehrkräfte sind im insgesamt sehr optimistisch, fast zwei Drittel der Noten sind gut oder sehr gut. Auf der Grundlage solch positiver Leistungsrückmeldungen können die Kinder kaum selbstkritischer und leistungsrealistischer werden.

Der bernische Lehrplan erwartet von der Leistungsbeurteilung einerseits, dass sie auf die Lernziele bezogen ist (absolute Leistungsnorm) und andererseits die Anstrengung des Kindes honoriert:

„Beurteilung dient in erster Linie der Analyse, Diagnose und Förderung des Lernens..... Die Beurteilung der Sachkompetenz hat primär lernzielorientiert zu erfolgen. Dabei soll aber auch der individuelle Fortschritt berücksichtigt und anerkannt werden“

(Erziehungsdirektion, 1995, Allgemeine Hinweise und Bestimmungen, 20).

Entsprechend wird seit 1995 auf Zeugnisnoten in den ersten beiden Schuljahren verzichtet, die Leistung wird lernzielbezogen beurteilt: „Lernziel erfüllt“, „teilweise erfüllt“ oder „übertroffen“, dies sind die Beurteilungskategorien des amtlichen Lernberichts. Zum Ende des dritten Schuljahres gibt es erstmalig Notenzeugnisse. Zusätzlich müssen sich die Kinder ab dritter Klasse selber einschätzen, also eine Selbstbeurteilung vornehmen, mit dem Ziel:

„Schülerinnen und Schüler sollen lernen, ihr eigenes Arbeiten zu beurteilen. Die Fremdbeurteilung kann in zunehmendem Masse durch die Selbstbeurteilung ergänzt und in gewissen Phasen auch ersetzt werden“ (Erziehungsdirektion, 1995, Allg. Hinweise und Bestimmungen, 20).

Der Lehrplan erlaubt also den Lehrkräften, die Leistungsbeurteilung positiver als objektiv gegeben, auszugestalten, sofern Kinder sich in hohem Masse anstrengen. Im Verlaufe der Schuljahre soll die Leistungsbeurteilung objektiver werden und zu einer realistischen

Einschätzung der eigenen Leistungsfähigkeit führen. Die Leistungsbeurteilung darf deshalb in den unteren Schuljahren insgesamt milder ausfallen als in späteren Schuljahren. Dies dürfte die Tendenz einer Klasse, sich im Durchschnitt zu überschätzen, fördern².

Die Pädagogik des Lehrplans fördert damit ein Beurteilungsverhalten, welches von der Befundlage der Psychologie gestützt wird. Für ein optimales Lernen ist eine leichte Überschätzung vermutlich günstig (Flammer, 2002). Taylor und Brown lösten mit ihren Aussagen, dass eine leichte Selbstüberschätzung für die Gesundheit und den Leistungserfolg am besten sei und deshalb gefördert werden müsse, eine Kontroverse aus (Block & Colvin, 1994; Taylor & Brown, 1994). Es ging um die Frage, ob nun eine leichte Überschätzung für den Menschen besser sei, oder eine möglichst realistische Selbsteinschätzung, wie es der „Common – sense“ unserer Gesellschaft erwartet. Am Ende der Kontroverse stellte die psychologische Forschung fest, dass eine leichte Überschätzung den besten Lernerfolg verspricht.

Allerdings ist der Preis dafür eine unrealistischere Selbstwahrnehmung. In Untersuchungen zur Kontrollmeinung, also zur Frage, ob jemand glaubt, ein bestimmtes Handeln erfolgreich ausüben zu können (aufgrund der Vorstellungen über die eigene Leistungsfähigkeit) fanden sich entsprechende Ergebnisse beispielsweise bei Little, Oettingen, Stetsenko und Baltes (1995). Sie untersuchten die Kontrollmeinung bei Kindern vom zweiten zum sechsten Schuljahr, die insgesamt etwas über 2000 Kinder umfasste. Die Befunde zeigen, dass amerikanische Schulkinder eine signifikant höhere durchschnittliche Ueberzeugung haben, ihren Schulerfolg selber kontrollieren zu können, als europäische. Gleichzeitig besteht bei diesen Kindern ein deutlich kleinerer Zusammenhang zwischen der Einschätzung der eigenen Leistungsfähigkeit und der tatsächlichen Leistungsfähigkeit (Korrelationen, je nach Alter, zwischen 0.16 und 0.32 für Los Angeles, für Westberlin 0.50 - 0.58): Die westberliner Kinder sind also präziser in ihren Selbsteinschätzungen. Oettingen, Lindenberger und Baltes (1992) hatten früher auch schon zwischen ost- und westberliner Kindern signifikante Unterschiede in der Beurteilung der Selbstwirksamkeit gefunden, die ebenfalls mit signifikanten Unterschieden in den Korrelationen von Einschätzung x Leistung einher ging: Die ostberliner Kinder schätzten sich signifikant weniger hoch ein, ihre Einschätzungen korrelierten schon ab der zweiten Klasse signifikant höher mit der Leistung als bei den Kindern aus Westberlin.

Die Bedeutung der Leistungsrückmeldungen durch die Lehrperson für die Höhe der Selbsteinschätzung der eigenen Fähigkeiten ist in den ersten Schuljahren gering. Wird sie durch die Kinder tatsächlich wahrgenommen (ab zweiter Klasse), dann stützt sie eher eine schon bestehende Überschätzung eigener Leistungsfähigkeit beim Kind, als dass sie zu einer

² Interessant ist diesbezüglich die aktuelle Diskussion (Winter 2004) um die Einführung der neuen Schülerbeurteilung: Einzelne Schülerkomites forderten eine Rückkehr zu Notenbeurteilungen.

Reduktion dieser Tendenz führt. Erst in den höheren Schuljahren der Grundschule sollen die Kinder dann zu realistischeren Vorstellungen eigener Leistungsfähigkeit finden. Entsprechend werden Leistungen mit zunehmendem Schulalter strenger beurteilt.

2.2.3 Determinanten der Schulklasse

Die Leistungsstärke der Mitschülerinnen und –schüler als komparativer Kontext hat einen bedeutsamen Einfluss auf das eigene Selbstbild (Jerusalem & Schwarzer, 1991). Rheinberg und Einstrup (1977) fanden beispielsweise unterschiedlich hohe Einschätzungen der eigenen Begabung bei gleicher Intelligenz in Abhängigkeit vom Klassenkontext: In Klassen, die einen relativ tieferen klassendurchschnittlichen Intelligenzgrad hatten, überschätzen sich die Kinder in ihrer Intelligenz stärker, als in Klassen mit höherem relativem Durchschnitt in der Intelligenz. Wenn also das klassenbezogene durchschnittliche Leistungsniveau im Verhältnis zur objektiven durchschnittlichen Leistung hoch ist, dann nimmt die Abweichung (in Richtung einer Überschätzungstendenz) ab. Die Höhe der Schätzabweichung in einer bestimmten Klasse ist also auch durch deren durchschnittliches Leistungsniveau im gemessenen Bereich beeinflusst. Für den Verlauf der Einschätzung über die Schuljahre hinweg ist das dann von Bedeutung, sollte sich das durchschnittliche Leistungsniveau der Klasse verändern. Das kann beim Übergang vom Kindergarten in die Schule der Fall sein: Wechselt ein Kind beispielsweise von einem mehrere Jahrgänge umfassenden Kindergarten in eine reine Jahrgangsklasse, dann wird die Leistungskonkurrenz zunehmen, weil es in der Schulklasse keine altersmäßig jüngeren Kinder mehr hat. Das durchschnittliche Leistungsniveau steigt. Die Schätzabweichung im Sinne einer Überschätzungstendenz müsste abnehmen.

Vom Kindergarten zur ersten Klasse sind Veränderungen im klassendurchschnittlichen Leistungsniveau leicht möglich: In der Untersuchung von Helmke (1991) wird das besonders deutlich. Nur 3 bis 4 Kinder aus den meist mehrere Jahrgänge umfassenden Kindergärten besuchten später zusammen die gleiche Schulklasse. Die anderen Kinder verblieben noch weitere Jahre im Kindergarten. Mit dem Übergang hat also der komparative Kontext gewechselt, die leistungsschwächeren Kinder bleiben noch im Kindergarten, die Bezugsgruppe der Erstklässler ist klassendurchschnittlich leistungsstärker geworden. Trotzdem wird die Selbsteinschätzung der Kinder in der ersten Klasse durch diese Determinante wenig beeinflusst. Wie wir noch sehen werden, orientieren sich die Kinder bei der Beurteilung ihrer Leistungsfähigkeit in diesen Jahren vorwiegend an ihrem eigenen Lernfortschritt. Die Leistungsstärke der Mitschülerinnen und –schüler bleibt weitgehend unbeachtet.

2.2.4 Determinanten des Individuums

2.2.4.1 Bisherige Leistung und persönlicher Standard

Die individuelle Lerngeschichte der Schulkinder ist in den ersten Jahren durch das „mastery-learning“ geprägt: Im Lernstoff wird erst weitergefahren, wenn das Bisherige verstanden und gelernt ist, wodurch auch bei potentiell schwächeren Kindern die individuelle Lerngeschichte eine von Erfolgen bestimmte ist. Vor allem das erste Schuljahr ist diesbezüglich bedeutsam, werden doch die absoluten „basics“ gelegt, ohne die es kaum ein angemessenes Weiterlernen gibt. Dies gilt vermutlich stärker für den Bereich Mathematik, als im Sport, wo nötigenfalls auch mit Stützunterricht nachgeholfen wird.

In beiden Fachbereichen gibt es in den ersten Schuljahren keinen Grund, die Leistungen selektiv zu beurteilen, denn erst am Ende des sechsten Schuljahres kommt es ja zu einem Übertritt in eine höhere Schule, respektive der Aufteilung in eine leistungsschwächere und eine –stärkere Gruppe von Kindern. Leistungsdifferenzen zwischen den Kindern aufzuzeigen ist in der Schuleingangsphase deshalb wenig wichtig, zentraler sind individuelle Erfolgserfahrungen auch für Schwächere. Selektivere Leistungsbeurteilungen und Leistungstests sind erst etwa ab dem vierten Schuljahr zu erwarten. Einen persönlichen Leistungsstandard, eine weitere Determinante des Leistungsselbstbildes, setzen sich Kinder schon mit 4 bis 5 Jahren (zusammenfassend: Oerter, 1995), also schon vor Eintritt in den Kindergarten: die einen Kinder sind eher erfolgsmotiviert. Sie stellen sich also bestimmten Herausforderungen gerne (beispielsweise schwierigen Mathematikaufgaben). Andere dagegen möchten Misserfolg unbedingt vermeiden und gehen kaum Risiken ein.

2.2.4.2 Der Wandel in der Bezugsnormorientierung

Das schweizerische Volksschulsystem ist unter anderem dadurch gekennzeichnet, dass es in den ersten fünf bis sechs Schuljahren alle Kinder eines Altersjahrganges integriert: Die Schülerinnen und Schüler werden unabhängig von ihren schulischen Leistungsmöglichkeiten miteinander die gleichen Klassen besuchen. Das Volksschulgesetz des Kantons Bern von 1992 hält in Artikel 17 fest, dass auch Kinder, deren Lernen durch Störungen und Behinderungen beeinträchtigt ist, in der Jahrgangsklasse unterrichtet werden sollen. Zusätzlich erhalten diese Kinder Spezialunterricht. Die Richtlinien (Erziehungsdirektion des Kantons Bern, 1997) sind so ausgestaltet, dass eine grösstmögliche Integration aller Leistungsstärken möglich ist. Über neunzig Prozent der Kinder dieser Altersstufe besuchen in der Schweiz

eine öffentliche Schule, die Kinder aller Leistungsniveaus integriert (Schweizerische Koordinationsstelle für Bildungsforschung, 1999). In 19 Kantonen erfolgt erst nach dem sechsten Schuljahr eine Aufteilung der Kinder nach dem Kriterium ihrer allgemeinen Schulleistungsfähigkeit, in fast allen anderen Kantonen erfolgt dieser Übergang nach dem fünften Schuljahr.

Bis zum Alter von 12 Jahren umfasst eine Schweizer Schulklasse Kinder fast aller Leistungsniveaus, was für das Fähigkeitsselbstbild nicht unproblematisch ist: Leistungsschwache Kinder sehen, was die Klassenbesten zu leisten im Stande sind.

Gesellschaftspolitisch verspricht man sich durch die Integration aller Jahrgangskinder in einer Volksschulklasse mehr Vor- als Nachteile, insbesondere einen Gewinn für den gesellschaftlichen Zusammenhalt. Eine vorzeitige Stigmatisierung der leistungsschwächeren Kinder durch deren Ausgliederung in Spezialklassen ist dabei ein anderes wichtiges Argument (Buff, 1991). Die eigene Schulklasse übt auf die Entwicklung des schulischen Fähigkeitsselbstkonzeptes einen starken Einfluss aus, an ihr entwickelt das Kind seine Vorstellungen eigener Leistungsfähigkeit (Geppert, 1997; Helmke, 1992). Schulkinder messen ihre Leistungen am Klassenmasstab. Ob sie sich leistungsmässig in einem Schulfach als „gut“ beurteilen oder nicht hängt nicht zuletzt vom Niveau der Vergleichsgruppe ab, dem entscheidenden Beurteilungsmasstab während der Grundschuljahre (Jerusalem, 1991). Erst im Jugendalter, am Ende der obligatorischen Schulzeit, verliert die Klasse als Bezugsgruppe zugunsten der Zugehörigkeit zu einem Schultypus an Einfluss auf das Selbstkonzept: Das Kind würde sich dann beispielsweise auf den Standpunkt stellen "besser eine schwache Sekundarschülerin, als eine gute Realschülerin" (Buff, 1991). Für die Schuleingangsstufe scheint aber diese interindividuelle Bezugsnorm noch wenig bedeutsam: Ruble et al. (1980) fanden, dass sich erst Zweitklässler von den Leistungen von Mitsreitern beeinflussen lassen, wenn es darum geht, eigene Leistungen zu werten. Die Bezugsnorm ist in Kindergarten und in der ersten Klasse vor allem der intraindividuelle Lernerfolg, also die Beurteilung, ob man nun mehr kann als beispielsweise vor einigen Wochen (Geppert, 1997, Oerter 1995). Noch jüngere Kinder im Vorschulalter berücksichtigen ebenfalls vorwiegend eigene Erfolgs- respektive Misserfolgserfahrungen, es bleibt dabei aber insbesondere die Berücksichtigung der Aufgabenschwierigkeit unbeachtet.

Kinder verändern im Verlaufe der Entwicklung den Bezugsmasstab, bis zum 6. Schuljahr nimmt die Orientierung am sozialen Masstab zu. Trotzdem bleibt der intraindividuelle Normbezug vorhanden. Das bedeutet, dass Kinder zunehmend aus mehreren Quellen Informationen über ihre Leistungsfähigkeit aufnehmen und nutzen, um sich ein Urteil zu bilden. Eine dritte Informationsquelle, die Kinder erst im Laufe der Grundschule in ihre Urteilsfindung integrieren sind die leistungsbezogenen Rückmeldungen der Lehrkraft. Aber auch wenn

Kinder die Leistungsrückmeldungen der Lehrkräfte bereits in den ersten Schuljahren stark berücksichtigen würden, würde dies an ihrer Beurteilung kaum etwas ändern, denn diese sind ja so positiv wie die eigene Erfahrung: „ Eigentlich kann ich die gestellten Aufgaben fast immer“. Befunde von Nicholls und Miller (1983) zeigen, dass Kinder erst im Alter von acht Jahren die Schwierigkeit einer Lernaufgabe daran einzuschätzen beginnen, wieviele ihrer Mitschüler dieser Aufgabe gewachsen sind. Vorher ist der zentrale Massstab der eigene Erfolg, oder allenfalls der Quervergleich mit leichteren oder schwereren Aufgaben. Bietet man Kindern konkret sichtbare Vergleichsmöglichkeiten an (Befunde von Heckhausen), dann benutzen sie diese Massstäbe zwar, leiten daraus aber kaum Schlussfolgerungen für die eigene Leistungsfähigkeit im allgemeinen ab (zusammenfassend: Oerter, 1995). Die eigene Leistungsfähigkeit kann erst realistisch eingeschätzt werden, wenn man über die Schwierigkeit einer Aufgabe im Bilde ist, das heisst, wenn man die Schwierigkeit einer Aufgabe unabhängig vom eigenen Können einschätzen kann. Diese Fähigkeit besitzen Kinder nach Befunden von Nicholls (1984) mit etwa 10 Jahren.

Zusammenfassend kann festgehalten werden, dass von der Schuleingangsphase bis in die fünfte, sechste Klasse Kinder in einer sehr leistungsheterogenen Lerngruppe ihre Schulzeit erfahren. Wenn in der Schuleingangsphase auch der eigene Leistungsfortschritt den stärksten Einfluss auf das Fähigkeitsselbstbild hat, so wird ab zweitem Schuljahr auch die eigene Schulklasse zum Bezugsmassstab. Weil das Leistungsgefälle in den Klassen gross ist, werden leistungsschwächere Kinder nun zunehmend merken, dass Andere deutlich besser sind. Das Fähigkeitsselbstbild gerät unter Druck. Die Beurteilung, ob man etwas gut oder sogar sehr gut kann, wird nach unten korrigiert. Die im Durchschnitt der Klasse ja sehr hohe Beurteilung eigener Fähigkeiten wird demgegenüber bei den leistungsstarken Kindern durch die zunehmende Bedeutung der Sozialnorm (als wesentlichem Bezugsmassstab) bestätigt. Insgesamt (starke und schwache Kinder zusammen) resultiert vermutlich eine Abnahme der durchschnittlichen Einschätzung, wegen der zunehmenden Bedeutung der sozialen Bezugsnorm dafür eine Zunahme der Schätzgenauigkeit.

2.2.4.3 Der Einfluss der Lernfreude im Schuleingangsalter

2.2.4.3.1 Der Verlauf der Lernfreude

Es ist im Schulalltag immer wieder ein eindrückliches Erlebnis zu sehen, wie Kinder mit Eifer und Interesse neuen Lernherausforderungen begegnen: Mit grosser Intensität und Ausdauer setzen sie hunderte von immer gleichen Buchstaben nebeneinander und rechnen mit Fingern einfache Additionen aus. Für Erwachsene faszinierend dabei ist, wie wenig sich diese Kinder von Misserfolgen oder von ihrer Langsamkeit beeindrucken lassen: sie sind ganz einfach stolz, ein bestimmtes Resultat erzielen zu können, und tragen ihre Zufriedenheit gerne auch zur Schau. Dieses innere Engagement und die Motivation tragen auch die Lehrkraft durch den Schulalltag. Es gehört dann leider zur schmerzlichen Erfahrung, feststellen zu müssen, wie diese Lernfreude Schritt für Schritt abnimmt und in der Oberstufe der Volksschule zu einer eigentlichen Lernunlust führt (zusammenfassend: Helmke, 1993).

Lehrkräfte erleben aber auch, dass es für die Kinder zum guten Ton gehört, Schule als etwas zu bezeichnen, das wenig Freude bereitet, während im Unterricht selbst die Kinder dann häufig sehr engagiert bei der Sache sind: Könnte es also sein, dass die zur Schau getragene allgemeine Unlust mit der Lernfreude in einzelnen Fachgebieten kontrastiert? Zur Lernfreude in einzelnen Bereichen gibt es für das Grundschulalter wenig empirisches Material, eine Untersuchung zu dieser Frage hat Helmke (1993) vorgestellt. Er hat die Entwicklung und den Verlauf der Lernfreude in den Fächern Deutsch und Mathematik vom Kindergarten bis zum 5. Grundschuljahr untersucht und die Ergebnisse für Knaben und Mädchen getrennt ausgewertet. Für beide Fächer gilt, dass die Lernfreude im 1. Schuljahr am höchsten ist, dann gegen Ende des 2. Schuljahres einen starken Abfall erlebt und anschliessend bis in die vierte Klasse mässig abnimmt. Im 5. Schuljahr, an dessen Beginn in Deutschland eine Selektion in verschiedene Schultypen stattfindet, ist nochmals ein Abfall der Lernfreude zu beobachten (Abb.2-4-1):

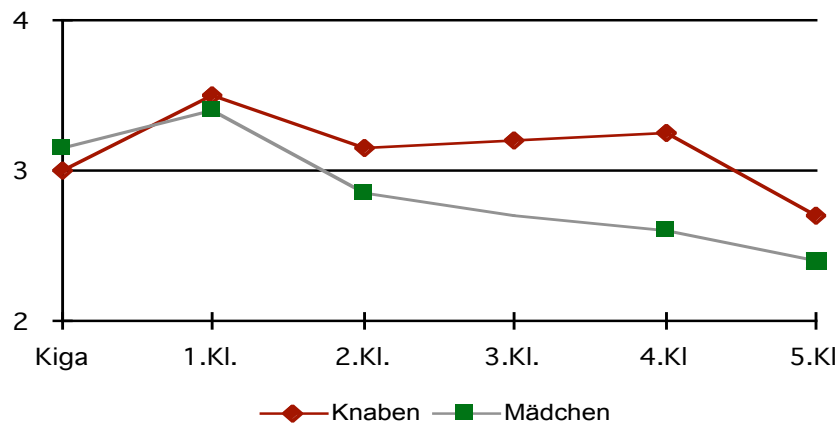


Abbildung 2-4-1 : Unterschiedliche Lernfreude von Knaben und Mädchen
 N= 114 Kinder, davon 58 Knaben; 2= mag ich manchmal, manchmal auch nicht;
 4= Mathematik mag ich sehr gern (aus: Helmke, 1997, S.70).

Der diskontinuierliche Verlauf der Lernfreude ist erklärungsbedürftig: Helmke (1991) vermutet, dass zu Beginn der Schulzeit eine Art Schonraum bestehe, die Lehrkraft würde dabei positives Feedback eher von der Anstrengung und dem Willen des Kindes abhängig machen und sich zu den gezeigten Leistungen noch gar nicht äussern. Falls dies zutrifft, hätte dies wahrscheinlich zur Folge, dass vor allem leistungsschwächere Kinder sich ihre Lernfreude nicht durch schwache Leistungen beeinträchtigen lassen. Helmke (1993) merkt an, dass noch andere Gründe zum Anstieg der Lernfreude vom Kindergarten zur ersten Klasse führen könnten: Eine Auswertung von entsprechenden offenen Fragen zur Lernfreude („warum magst Du Mathematik gut, oder nicht so gut?“) ergibt nach Helmke die Befunde, die wir im vorderen Kapitel gefunden haben. Die Lernfreude ist durch tatsächliche Leistungen im Kindergarten noch wenig beeinflusst und auch der soziale Vergleich scheint kaum auf die Lernfreude zu wirken. Erst gegen Ende der zweiten Klasse würden entsprechende Einflüsse in den Antworten der Befragung sichtbar.

Die Anwendung der sozialen Bezugsnorm beeinflusst die Lernfreude erst in höheren Schuljahren (z. B. Buff, 1991; Jerusalem & Schwarzer, 1991; Stipek & Mc Iver, 1989). In Abb. 2-4-1 zeigt sich, dass die Lernfreude über den gesamten Zeitraum erhalten bleibt. Knaben und Mädchen sind in Mathematik, aber auch in Deutsch positiv eingestellt. Die Verlaufsgestalt ist in beiden Fächern ähnlich, in Mathematik kann ein geschlechtsspezifischer Unterschied in der Lernfreude zu Gunsten der Knaben festgestellt werden, der ab drittem Schuljahr schwach signifikant wird. Auffällig ist der stärkere Abfall der Lernfreude der Knaben nach dem Übergang in die Sekundarschule: Die geschlechtstypisch höheren Erwartungen der

Knaben an ihre Leistungsfähigkeit in Mathematik werden in der 5. Klasse vermutlich von der Realität eingeholt.

Können wir also, wie eingangs behauptet, von einem Abfall der Lernfreude während der Schuleingangsphase sprechen? Die Lernfreude des Schulanfängers sinkt zwar vom Kindergarten bis in die fünfte Klasse ab, ist aber bei den Knaben, bezüglich des Lernens von Mathematik am Ende der vierten Klasse immer noch höher als im Schuleintrittszeitraum. Für die Mädchen hingegen ist eine langsam fallende Lernfreude festzustellen, die aber insgesamt immer noch positiv ist. Die Lernfreude nimmt also nur bei den Mädchen ab, während sie bei den Knaben über Jahre praktisch stabil bleibt, ja ganz leicht zunimmt. Der starke Abfall bei den Mädchen führt dazu, dass im insgesamt der Kinder trotzdem ein leichter Abfall der Lernfreude festzustellen ist (Helmke, 1997). Die Befunde von Helmke bezüglich der Entwicklung der Lernfreude sind die einzig verfügbaren für Kinder im frühen Grundschulalter. Für die höheren Schuljahre gibt es einige Befunde: Pekrun und Fend (1991) haben in einem Überblicksartikel Daten zur Sekundarstufe (ab 5. Schuljahr) zusammengefasst, welche eine Abnahme der Schulfreude bis hin zur Schulunlust feststellen. Flammer und Alsaker (2002) zitieren Befunde, die darauf hinweisen, dass generell während der Adoleszenz die Schulunlust zunimmt.

Als Determinante des Fähigkeitsselbstbildes wirkt sich die Lernfreude in Mathematik positiv aus. In der ersten Klasse ist sie am höchsten und verläuft ähnlich wie die Einschätzungshöhe. Am Ende der zweiten Klasse ist ein leichter Abfall der Lernfreude messbar, der bei den Mädchen signifikant ist. Das deutet auf die Möglichkeit hin, dass die Lernfreude als Determinante in der zweiten Klasse zu einem Geschlechtseffekt führen könnte: Bei Mädchen könnte ein stärkerer Abfall der Selbsteinschätzung in Mathematik die Folge sein.

2.2.4.3.2 Lernfreude und Leistung beeinflussen sich gegenseitig

Die Lernfreude des Kindes ist ein Indikator für die emotionale Bewertung eines täglichen schulischen Lernfeldes, sie hat Einfluss auf Motivation und Leistungsfähigkeit. Die Lernfreude als affektive Komponente der Einstellung zum Lernen ist eine der beiden Komponenten vieler Motivationstheorien (z. B. Heckhausen, 1980), nämlich die Valenzkomponente. Neben diesem Wertaspekt steht als zweite Komponente der Erwartungsaspekt im Zentrum, also die Frage danach, ob sich ein Kind eine bestimmte Leistung, die es eigentlich erbringen möchte ("Ich habe Freude an Mathematik") auch tatsächlich zutraut ("Ich kann das"). Während Motivationstheorien sich für die Ausprägung von Valenz und Kompetenz und deren Zusammenspiel bezüglich der Gesamtmotivation interessieren, steht hier vor allem die Enge des

Zusammenhangs von Valenz und Kompetenz im Vordergrund: Setzt eine hohe Valenz im Bezug auf ein Schulfach auch eine hohe Kompetenz voraus? Pädagogisch ausgedrückt: Haben diejenigen Kinder an Mathematik Freude, die glauben, dass sie in diesem Fach sehr leistungsfähig sind? Haben andererseits diejenigen Kinder keine Freude an Mathematik, die sich als eher leistungsschwach einstufen? Helmke (1993) hat folgende Korrelationen zwischen Lernfreude und Mathematikleistung gefunden (Tab.2-2-3):

Tabelle 2-2-3: Korrelation von Lernfreude und Mathematikleistung (Helmke, 1993)

KG	1.Kl.	2.Kl.	3.Kl.	4.Kl.
0.33	0.35	0.31	0.63	0.67

Die Überprüfung der Zusammenhangsstärke zwischen der Lernfreude und der Leistung im darauffolgenden Jahr ergibt bei Helmke, dass die Lernfreude bedeutend stärker von der Leistung abhängig ist als umgekehrt, pädagogisch ausgedrückt: die zunehmende Leistungsstärke erhöht die Freude am Fach im folgenden Jahr ($r=0.49$), während die Freude an sich sehr schwach mit der Leistung des folgenden Jahres korreliert ($r=0.16$).

In einer weiteren Studie mit 1113 Kindern und einem Messintervall von 6 Monaten hat Helmke (1997) folgende Korrelationen zwischen Lernfreude und Mathematiknote gefunden (Zeitspanne 2. bis 4. Klasse):

Tabelle 2-2-4: Korrelationen von Lernfreude und Mathematiknote (Helmke, 1997)

2.Kl. I	2.Kl. II	3.Kl. I	3.Kl. II	4.Kl.
0.25	0.26	0.32	0.32	0.35

Diese Befunde von Helmke entsprechen ungefähr denjenigen aus den siebziger Jahren, also einem schwachen Zusammenhang von Leistung und Lernfreude, wie sie Aiken (1976) in einem Überblicksartikel referiert. Es zeigt sich, dass die Lernfreude und die objektiv gemessene Leistung in der Schuleingangsphase nur schwach korrelieren, der Zusammenhang wird kaum stärker. Die effektive Leistungsfähigkeit ist zur Erklärung der Lernfreude vom Kindergarten bis zur zweiten Klasse kaum bedeutsam, auch leistungsschwächere Kinder haben also eine hohe Lernfreude. Der diskontinuierliche Verlauf der Lernfreude ist als Determinante der Performanz relativ unbedeutend.

2.2.4.4 Der Aufbau und die Ausdifferenzierung des Attributionskonzeptes

Ein grosser Teil der Determinanten des Leistungsselbstbildes ist während der Schuleingangsphase kaum Veränderungen unterworfen. Anders ist dies bei der Determinante „Attribuierungen“: In der Schule erlebt das Kind viele Lernerfolge, aber auch Misserfolge. Die Art und Weise wie es den Erfolg, respektive Misserfolg erklärt nimmt zentral Einfluss auf die Ausgestaltung des Leistungsselbstbildes.

Sensibel auf Erfolge / Misserfolge reagieren schon kleine Kinder: Bereits mit zweieinhalb Jahren können sie ein Produkt ihres Handelns auf sich selbst als Verursachende zurückführen (Heckhausen, 1985) und schon mit drei Jahren reagieren sie mit Stolz oder Beschämung, wenn sie erfolgreich, respektive misserfolgsreich sind (Geppert & Heckhausen, 1990). Kinder wissen also früh, was sie können und was nicht. Sie entwickeln ihr Wissen über die eigene Handlungskompetenz schon vor der Schulzeit. Aus dieser Erfahrung aber folgerichtig ableiten zu können, ob sie selber nun in einem Leistungsbereich „gut“ oder „schwach“ sind, also (realistische) Beurteilungen über die Qualität eigener Handlungskompetenz zu entwickeln, dauert viel länger: Ruble, Parsons und Ross (1976) legen den Zeitpunkt für die Möglichkeit des Kindes, diesbezüglich realistische Urteile zu kommunizieren, frühestens auf das Alter von 7-9 Jahren fest. Nicholls (1978, 1983) kommt zum Schluss, dass erst nach dem zehnten Lebensjahr bei den Kindern eine angemessene Vorstellung über das Zusammenwirken verschiedener Ursachen von Erfolg / Misserfolg erwartet werden kann. Also erst zu Beginn der vierten Klasse wären dann die Kinder allgemein in der Lage, das Ursachenzuschreibungsmodell richtig zu verstehen. Der Aufbau der entsprechenden Kognitionen entwickelt sich nach Flammer (1990, 1995, 1996) über verschiedene Stufen, die in der folgenden Abbildung zusammenfassend dargestellt sind:

Stufe	Struktur	Bezeichnung des kognitiven Elements	Alter
0	Effekte ereignen sich	-----	ab Geburt
1	Erfahrungen von Effekten	Ereignisschema : externe Ursache - externe Wirkung kennen	ab 4 Monaten
2	Selber Effekte erzeugen	Kausalschema: interne Ursache- externe Wirkung kennen	
3	Wissen,dass man selber bestimmte Effekte erzeugt	Internes und externes Kausalschema verstehen	ab 18 Monaten
4	Selbstbewusst wissen, dass man Effekte nach Erfolgskriterien erzeugt hat	Internes Kausalschema mit Qualitätsschema kombinieren	ab 3. Jahr
5	Selbstbewusstsein genereller Effektkontrolle	Kontrollschema, resp. Kontrollmeinung	vor KG
6	Die Bedeutung der Anstrengung für Leistungserfolg erkennen	Die Dimension stabil – intern labil intern im Ursache – Wirkungsschema ausdifferenzieren	ab erster Klasse
7	Verstehen wie Aufgabenschwierigkeit und Fähigkeit zusammenwirken	Die Dimension stabil intern-stabil extern als Ursache-Wirkungsschema verstehen	ab dritte Klasse
8	Die kompensatorische Wirkung von Anstrengung und Fähigkeit verstehen	Sämtliche Kombinationsmöglichkeiten des Attributions-schemas flexibel und reversibel verstehen	ab vierter Klasse

Abbildung 2-2-2: Aufbau der Kognitionen zur Bildung angemessener Attributionen (nach Flammer, 1996)

Generalisierte Vorstellungen einer Ursachenzuschreibung verfestigen sich etwa im Alter von fünf Jahren (Stufe 5): Erfolgreiches Handeln wird ursächlich mit dem eigenen Tun in Verbindung gebracht. Dieses Wissen bezieht sich dabei auf Handlungsfelder, in denen Kinder über ausreichende Erfahrungen verfügen, beispielsweise im Bereich Bewegung/Sport.

Bei Beginn der Schuleingangsphase befindet sich das Kind auf dieser Stufe. Das ist sein erster Tüchtigkeitsmassstab. Nach dem Kindergarten, im Verlauf der ersten Klasse, beginnen sich die kognitiven Vorstellungen der Bedeutung der Anstrengung für den Lernerfolg auszu-differenzieren (Stufe 6), die Anstrengung wird zur Hauptkovariate des Erfolgs (Flammer, 1990). An Einfluss gewinnt zunehmend die Aufgabenschwierigkeit (Skinner, Chapman & Baltes, 1988). Erst zu Beginn der dritten Klasse verstehen die Kinder dann, dass für die Bewältigung von schwierigen Aufgaben mehr Fähigkeiten nötig sind und diese nicht beliebig durch Anstrengung kompensiert werden können (Stufe 7), sie sehen also, dass andere trotz Anstrengung weniger fähig sind. Jetzt lernt das Kind zu unterscheiden zwischen dem zeit-stabilen Fähigkeits- und dem zeitvariablen Anstrengungskonzept. Wichtig ist, dieses nicht nur intuitiv zu verstehen, sondern dieses auch verbalisieren zu können (Geppert, 1997). Noch verstehen die Kinder aber das kompensatorische Verhältnis von Anstrengung und Fähigkeit nicht richtig, das gelingt erst später (Kun, 1977; Stipek & De Cotis, 1988). Ob nun die eigenen Fähigkeiten, die Anstrengung oder die Aufgabenschwierigkeit einen entscheidenden Anteil am Leistungserfolg haben ist ja auch für Erwachsene nicht immer eindeutig zu bestimmen (Flammer, 1990), diese Aspekte aber überhaupt richtig zu unterscheiden und zu verstehen, gelingt den Kindern erst am Ende des ersten Lebensjahrzehnts.

2.2.5 Zusammenfassung

Im euro-amerikanischen Raum nimmt die Bedeutung der Selbstverwirklichung für das Wohlbefinden des Menschen zu. Vorstellungen über eigene Stärken und Schwächen werden den Menschen bewusster und das Bedürfnis, das eigene Leben selber zu steuern wichtiger. Diese Entwicklung hat auch vor der Schule nicht halt gemacht, Sie hat die Lehrpläne, Lehrmittel und Unterrichtsdidaktik in den letzten zwanzig Jahren zunehmend beeinflusst und zur Forderung geführt, dass Kinder in ihrer Selbständigkeit gefördert werden sollen. Auch das Selbst rückt zunehmend ins Zentrum, schon ab der dritten Klasse sind Kinder angehalten, sich selbst bezüglich Sach-, Selbst- und Sozialkompetenz verbindlich einzuschätzen.

Die Förderung des Wissens um eigene Stärken und Schwächen ist allerdings für leistungsschwächere Kinder schmerzhaft, sobald sie merken, dass Mitschüler und Mitschülerinnen deutlich besser sind. Dem versucht die Schule entgegenzuwirken, indem die Leistungsbeurteilung stärker individualisiert wird durch die Mitberücksichtigung des Lernfortschritts. Leistungsbeurteilung soll zudem den Fleiss und Willen des Kindes belohnen, insbesondere in den ersten Schuljahren. Kinder erhalten somit Leistungsfeedback, das zur Überschätzung der eigenen Leistung führen kann. Psychologische Befunde zeigen, dass Kinder

im Kindergarten und in den ersten beiden Schuljahren aber vor Allem den eigenen Lernfortschritt zum Massstab der Einschätzung eigener Leistungsfähigkeit machen. Ab der zweiten Klasse berücksichtigen sie zunehmend auch die soziale Norm und die Rückmeldungen der Lehrperson. Da im Kindergarten und den unteren Jahren der Volksschule eine sehr grosse Leistungsheterogenität besteht, werden die leistungschwächeren Kinder zunehmend realisieren, dass die meisten anderen Kinder leistungsstärker sind. Vom dritten Schuljahr an (in Deutschland schon ab zweitem Schuljahr) kommen Leistungsbeurteilungen in Form von Noten / Zensuren dazu, die in den oberen Grundschuljahren selektiv werden. Insgesamt führt dies dazu, dass Kinder eine zunehmend realistischere Vorstellung eigener Leistungsfähigkeit gewinnen und dass die Überschätzung der Kinder zwischen dem Kindergartenalter und den oberen Grundschuljahren signifikant abnimmt. Für den Kindergarten und die beiden ersten Schuljahre gilt aber generell die pädagogische Grundhaltung des „Wohllollens“, mit der Folge, dass Kinder eher zur Überschätzung neigen. Die absolute Einschätzhöhe dürfte wegen des grossen Lernfortschritts, den die Kinder im ersten Schuljahr bei sich wahrnehmen, kaum abnehmen, denn sie beurteilen ihre Leistungsfähigkeit ja vorwiegend am individuellen Lernzuwachs. Die Genauigkeit der Einschätzung nimmt zu (Korrelation), weil Kinder zunehmend ihre klassenrelative Rangposition erkennen. Die Unterscheidung der Entwicklung der Höhe der absoluten Selbsteinschätzung von der zu erwartenden Zunahme der klassenrelativen Einschätzgenauigkeit (accuracy) ist deshalb wichtig.

Im Schuleingangsbereich, der ja für diese Entwicklungsfragen besonders interessant ist, weil in diesem Alter wichtige gesellschaftliche Fertigkeiten wie Lesen, Schreiben und Rechnen erstmals im Leben unter typischen Wettbewerbsbedingungen und unter gleichaltrigen Menschen erlernt werden, ist das Leistungsspektrum der Kinder so breit, wie später im Leben kaum mehr: die Chance, sich mit dem Durchschnitt zu messen ist hier am höchsten. Aber erst im zweiten Lebensjahrzehnt ist das Kind von seinem Denken her in der Lage, sich angemessen einzuschätzen, also etwa in der vierten Klasse.

Im nächsten Kapitel geht es um die Frage, in welchen Situationen das Fähigkeitsselbstkonzept in offenen Lernsituationen von Bedeutung ist und wie sich die Schätzgenauigkeit auf das Aufgabenwahlverhalten auswirkt

2.3. Befunde zur Entwicklung und Ausdifferenzierung spezifischer Kausalitätskonzepte

2.3.1 Anstrengung und Fähigkeit als Ursache von Erfolg unterscheiden lernen

Nicholls (1978) hat Kindern diverse Situationen von kindlichem Leistungsverhalten in Form von Filmen vorgespielt. Nachher mussten die Kinder das jeweilige Leistungsversagen, respektive den –erfolg der beobachteten Kinder erklären. Diese Erklärungen wurden aufgenommen und anschliessend analysiert. Nicholls fand vier Niveaus von Erklärungen:

Niveau 1:

Anstrengung und Fähigkeit werden nicht systematisch als voneinander unabhängig betrachtet und deshalb auch vermischt.

Niveau 2:

Anstrengung wird als Hauptursache von Leistung bezeichnet. Der Einfluss der Fähigkeit auf ein Ergebnis wird noch nicht richtig verstanden.

Niveau 3:

Dass Fähigkeit neben der Anstrengung auch eine Rolle spielt, wird auf diesem Niveau in Erwägung gezogen, wenn beispielsweise das Ergebnis einer Leistung gleich ist, bei offensichtlich unterschiedlicher Anstrengung.

Niveau 4:

Nun sind Kinder in der Lage, bei allen Leistungsbeobachtungen systematisch und ohne Einschränkungen Anstrengung und Fähigkeit kompensatorisch als Ursache von Leistungshandeln zu verstehen. Sie wissen auch, dass die Fähigkeit unabhängig von der Anstrengung ist. Kinder verschiedenen Alters befinden sich nach diesen Befunden von Nicholls in folgenden Niveaus (Tab.2-3-1):

Tabelle 2-3-1: Anzahl Kinder, die auf diesem Niveau geantwortet haben
Total 16 Kinder pro Jahrgang,
Durchschnittsalter der Kinder in Jahren (Nicholls 1978, S.805)

	5J	6J	7J	8J	9J	10J	11J	12J	13J
N 1	16	11	3	2	1	0	0	0	0
N 2	0	5	10	10	9	4	6	2	0
N 3	0	0	3	3	5	7	5	5	4
N 4	0	0	0	1	1	5	5	9	12

Zwei Drittel der 7 - 9 jährigen Kinder sind auf dem Niveau 2: Leistung ist in ihrer Vorstellung vor allem von Anstrengung abhängig. Die Entwicklung in diesem Altersabschnitt ist eher klein. Die 10 – 11 jährigen Kinder sind überwiegend im Niveau 3 und 4. Im Alter von 12 Jahren, also im sechsten Schuljahr, argumentieren die Kinder fast alle auf Niveau 3 und 4.

Für die Schuleingangsphase gilt für die Kinder die Grundüberzeugung „Wer sich anstrengt, ist gut“. Die Kinder befinden sich in Stufe I der Selbstkonzeptentwicklung gemäss dem Modell von Harter (1983). Sich anstrengen bedeutet hier, dass man in einer konkreten Leistungssituation handeln soll, wenn man erfolgreich sein will. Befunde von Pfister (1992) unterstützen diese Annahmen: Sie befragte Zweitklässler danach, wie sie Leistungserfolg / -misserfolg erklären. Bei Leistungsaufgaben zum Sport (eine bestimmte Strecke schwimmen können) und zur Mathematik nannten sie vorwiegend zwei Arten von Begründungen: Entweder Aussagen wie „ich kann das einfach“ oder „ich kann diese Aufgabe, weil ich es schon erfolgreich versucht habe“. Auf das konkrete Können rekurrierende Begründungen machen sowohl bei der Sport- wie bei der Rechenaufgabe 78% aller Antworten aus.

Ein weiterer Befund, der in die ähnliche Richtung geht (Wicki, Reber, Flammer, Grob, 1994) zeigt, dass Kinder mit zunehmendem Alter verschiedenartigere Begründungen für Erfolg / Misserfolg referieren. In dieser Untersuchung nannten die 7 - 9 jährigen Kinder (1.-3. Klasse) als Hauptgrund für Erfolg / Misserfolg das eigene Handeln (31% der Begründungsantworten) oder die Schwierigkeit respektive Leichtigkeit der Aufgabe (28% der Antworten). In höherem Alter nahmen dann Begründungen mit Hinweis auf die eigenen Fähigkeiten und auf die eigene Lernlust / Motivation deutlich zu. Bei den Leistungssituationen ging es in dieser Untersuchung vorwiegend um Fragen der Selbst- und Sozialkompetenzeinschätzung. Immerhin zeigen aber die Befunde in die gleiche Richtung, wie bei Pfister.

2.3.2 Die Beurteilung der Aufgabenschwierigkeit

Nicholls & Miller (1983) untersuchten, an welchen Massstäben Kinder die Schwierigkeit einer Aufgabe messen und kamen zu einem dreistufigen Ablauf:

Stufe 1:

Kinder haben eine selbstzentrierte Vorstellung der Aufgabenschwierigkeit, sie messen die Schwierigkeit der Aufgabe daran, ob sie sie selber lösen können, respektive ob sie glauben, diese lösen zu können. Auf dieser Stufe befinden sich allerdings nur noch wenige Erst- (12%) und Zweitklässler (2%) (Nicholls & Miller, 1983).

Stufe 2:

Kinder haben eine objektive Vorstellung der Aufgabenschwierigkeit, sie können sichtbar unterschiedliche Aufgabenschwierigkeiten erkennen, um die Schwierigkeit einer Aufgabe zu beurteilen. Voraussetzung ist, dass sie die unterschiedlichen Schwierigkeitsgrade konkret sehen können, um die Unterschiede zu beurteilen. Die meisten Erstklässler (60%) sind auf dieser Stufe (Nicholls & Miller, 1983).

Stufe 3:

Kinder orientieren sich bezüglich Aufgabenschwierigkeitseinschätzung an den Erfolgen / Misserfolgen von Mitschülerinnen und – schüler. Auf Stufe 3 befinden sich von den Zweitklässlern bereits 55%.

Eine Schwierigkeit bei der Übertragung der Befunde von Nicholls & Miller (1983) ist, dass die Kinder unserer Untersuchung durchschnittlich fast ein Jahr älter sind. Dies dürfte sich so auswirken, dass mehr als 55% der Zweitklässler unserer Untersuchung soziale Quervergleiche machen um die Aufgabenschwierigkeit einzuschätzen. Die Aufgabenschwierigkeit als Kriterium für Erfolgs- / Misserfolgseinschätzungen ist allerdings in der Schuleingangsphase noch wenig von Bedeutung, Hauptkriterium bleibt die eigene Anstrengung.

Fazit: Während der Schuleingangsphase zeigen die Befunde keine wesentlichen Veränderungen in Bezug auf die Art der Ursachenbegründungen, wenn nun beispielsweise die Anstrengungsbereitschaft konstant hoch bleibt, die ja auch in den Augen der Lehrkraft ein wichtiges Kriterium der Leistungsbeurteilung darstellt, dann sind konstant hohe Selbsteinschätzungen durch die Kinder zu erwarten: „Ich strenge mich immer an, also bin ich auch gut, respektive erfolgreich.“ Erst in späteren Jahren wird die Lehrkraft vermehrt lernzielorientiert beurteilen und ein grösserer Teil der Kinder wird bei sich selbst feststellen, dass auch bei hoher Anstrengung nur mittlere Leistungen möglich sind. Bezogen auf eine Schulklasse führt dies dannzumal zu einem Absinken der durchschnittlichen Leistungseinschätzung.

2.3.3 Wirkungen der Ausdifferenzierung von Fähigkeit und Anstrengung auf Höhe und Genauigkeit der Selbsteinschätzung

Für den Zeitraum vom letzten Quartal im Kindergarten bis zum Ende des zweiten Schuljahres tendieren Kinder dazu, Anstrengung für den wichtigsten Grund für Erfolg im schulischen Leistungsbereich zu halten. Sie können aber bereits andere Gründe anführen, um Leistungserfolg zu erklären, insbesondere den Einfluss der Aufgabenschwierigkeit. Das kompensatorische Verhältnis von Anstrengung und Fähigkeit und insbesondere die Unabhängigkeit der Fähigkeit von der Anstrengung ist ihnen noch nicht vertraut. Gegen Ende der Schuleingangsphase verschiebt sich aber die Beurteilung der Aufgabenschwierigkeit zunehmend vom eigenen Leistungserfolgsmasstab zu demjenigen der Gruppe: was allen schwer fällt, muss auch schwierig sein, es gilt aber auch, dass was man selbst nicht kann, nicht unbedingt schwierig sein muss. Während des Schuleingangsalters wird also die Bezugsgruppe zunehmend wichtiger, wenn es gilt, Aufgaben als schwierig oder einfach zu taxieren. In einer leistungsstarken Klasse werden also schwierige Aufgaben als leichter empfunden, eigenes Versagen deshalb eher mit mangelnder Anstrengung oder Fähigkeit erklärt, als in einer leistungsschwachen Klasse.

Während Befunde zeigen, dass Kinder erst im zweiten Schuljahr die Leistungen der anderen Kinder zum Masstab eigener Leistungsfähigkeit machen, so beurteilen sie doch die Aufgabenschwierigkeit und damit indirekt die eigene Leistungsfähigkeit an der Klassenleistungsfähigkeit. Bis zum Ende der zweiten Klasse können Kinder noch kaum beurteilen, ob ihre Leistungen nun an einem absoluten Masstab gemessen „gut“ oder „schlecht“ oder „genügend“ sind, sie verlassen sich hier auf die Leistungen der Klasse. Weil in den ersten Schuljahren alle Kinder das Gefühl haben, sie seien gut (Normbezug ist der eigene Leistungsfortschritt) und später merken, dass sie auch von der Lehrkraft entsprechendes Feedback erhalten, werden die Kinder weniger genau beurteilen können, wo sie in einer Klassenrangliste denn stehen. Dies gelingt besser am Ende der zweiten Klasse, weil nun die meisten Kinder die soziale Bezugsgruppe als Masstab eigener Leistungsfähigkeit teilweise einbeziehen. Die Korrelation zwischen Rangplatzeinschätzung und tatsächlichem Rangplatz nimmt zu. Solange die Kinder glauben, dass die Anstrengung dafür verantwortlich ist, ob man Erfolg hat oder nicht, ist ihre Lernmotivation hoch. Zudem ist das Gefühl vorherrschend, man sei gut (wer sich anstrengt, ist gut, wer faulenz, schlecht). Alle diese Faktoren tragen dazu bei, dass sich Kinder im Klassendurchschnitt überschätzen. Erst wenn die soziale Bezugsnorm wichtig wird, wird dieser Durchschnitt sinken. Voraussetzung ist, dass die leistungsschwächsten Reif genug sind, die soziale Norm zu integrieren. Dies geschieht für diese Gruppe eher spät,

was dazu führt, dass bis zum Ende der zweiten Klasse die klassendurchschnittliche Einschätzung vermutlich konstant hoch bleiben wird.

2.4 Empirische Befunde zur Verlaufsgestalt der Selbsteinschätzung im Kindergarten- und Grundschulalter

Die Befunde zum Verlauf der Einschätzhöhe in der Schuleingangsphase zeigen, dass die Selbsteinschätzung eigener Fähigkeiten noch sehr hoch ist und in den ersten Jahren erst wenig abnimmt (Eshel & Klein, 1981; Nicholls, 1978; Stipek & Daniels, 1988; Stipek & Tannatt, 1984). Bei einem Vergleich der Befunde von Erstklässlern mit Kindern des vierten / fünften Schuljahres ist die Abnahme aber signifikant. Eine genauere Betrachtungsweise des Verlaufes in den ersten Jahren, insbesondere unter Einbezug des Kindergartens, zeigt, wie schon dargestellt, dass dieser Entwicklungsverlauf weder linear noch durchgängig abnehmend ist (Helmke, 1991, 1999). Schon Stipek und Mc Iver (1989) hielten in ihrem Überblicksartikel fest, dass eine Abnahme der Überschätzung nur für das generelle Fähigkeitskonzept („Wie gut bist du in der Schule?“) zutrefte, befrage man dagegen konkrete Leistungsbereiche (Beispielsweise: „Wie gut kannst du rechnen?“), sei die Abnahme in den ersten Jahren sehr gering. Die Befunde dieser Studien werden nun dargestellt und diskutiert (Tab.2-4-1). Der Übersichtlichkeit halber werden im folgenden zuerst die oben schon zitierten Untersuchungen tabellarisch dargestellt, anschliessend erläutert und diskutiert. Daran folgen die Untersuchungen aus den 90-er Jahren (Tab.2-4-2) und anschliessend neuste Befunde (Tab.2-4-3), die nach der Erhebung der Kindergartendaten der vorliegenden Arbeit erhoben oder publiziert wurden.

In der folgenden Tabelle 2-4-1 finden sich die Rohwerte der jeweiligen Skala. Mit Ausnahme der ersten Untersuchung bedeuten die sinkenden Mittelwerte (M) auch eine Abnahme der Einschätzhöhe. Der zweite Wert stellt den Korrelationskoeffizienten dar, also die Einschätzgenauigkeit (accuracy). Einbezogen wurden Studien, welche Daten über Kinder im Kinderartenalter beinhalten (Ausnahme: Eshel & Klein, 1981). Auf der rechten Seite ist dargestellt, auf welchen Fachbereich bezogen die Kinder ihre Leistung zu beurteilen hatten und ob sie dies in Bezug zur objektiven Leistung (absolute Norm) zu beurteilen hatten, oder in Bezug zur Klassenrangreihe (soziale Norm).

Tabelle 2-4-1: Mittelwerte und Korrelationen von Leistungseinschätzungen (1978–1989)
(sehr niedrige r = „0“; keine Angaben zu r „-“)

Autor, Stichproben- Beschreibung		Vor- KG	KG	1.KI	2.KI	3.KI	4.KI	5.KI	6.KI	Beurteilungsbereich Normbezug, Skala Beispiel
Nicholls, 1978 N=8-10 pro Jahrgang	M r	3.06 0	5.12 0	9.06 0.21	9.00 0.27	11.9 0.58	13.8 0.71	11.6 0.57	12.9 0.80	Wie gut kann ich lesen? Soziale Norm, Skala : 1-30 Ich bin das beste Kind =01 Ich bin das schwächste K. =30
Eshel, Klein 1981 N=2199 obere Zahl: Unterschicht untere Zahl: Mittelschicht	M M r r			4.75 4.76 0.09 0.01	4.57 4.73 0.21 0.26	4.39 4.51 0.29 0.41	4.05 4.38 0.39 0.63			Wie gut kannst du rechnen ? Absolute Norm, Skala: 1-5 Ich kann sehr gut rechnen =5 Ich kann nicht gut rechnen =1
Stipek & Tannatt, 1984, N=32 pro Gruppe: Grp. 1=VK Grp. 2=KG, 1.KI. Grp. 3=2.KI, 3.KI.	M r	4.9	4.7	4.7	4.2	4.2				Wieviele Sterne bekommst du für deine Schulleistungen ? Absolute Norm, Skala : 1-5 Ich bin sehr gut, erhalte 5 Sterne
Beneson & Dweck , 1986 N=144, total	M r		4.7	4.9	4.4		4.0			Wieviele Sterne bekommst du für deine Schulleistungen ? Absolute Norm, Skala 1-5 Ich bin sehr gut, erhalte 5 Sterne
Frey, Ruble, 1987 N= 83, total	M r		2.52	2.22	2.01		1.56			Lesefähigkeit , Soziale Norm, Skala -20 bis +20 Abweichung von Lehrperson- Urteil: 2 =2 Ränge höher einge- schätzt als Lehrperson
Stipek , Daniels, 1988 ,N=80 total oben : viel unten: wenig Leistungsfeed- back	M M		15.1				14.8			Allgem. Schulkompetenz: Wie gut bist du , im Vergleich zu den Anderen? Soziale Norm, Skala 1–20, ich bin die Beste =20

Die Studie von Nicholls (1978) beruht auf einer sehr niedrigen Anzahl Probanden und erfasst das Leistungsselbstbild im Bereich „Lesen“, also einer Tätigkeit, die erst mit dem Schuleintritt bedeutsam und lernzielorientiert gefördert wird. Das Kind hatte zu beurteilen, an welcher Position es in der Leistungsrangreihe steht, es standen 30 Positionen zur Verfügung. Die Einschätzposition nimmt vom Kindergarten zur ersten Klasse im Lesen ab und bleibt in der zweiten Klasse stabil. Vermutlich resultiert dieser Befund aus der Tatsache, dass die Kinder im Kindergarten ihre Leistungen im Lesen noch gar nicht vergleichen konnten, weil Lesen nicht unterrichtet und gelernt wird. Mangelnde Vergleichsmöglichkeiten wären dann die Ursache für die Befundlage. Nach der zweiten Klasse sinkt die Einschätzposition weiter ab (die Zahl bedeutet den geschätzten Rangplatz, „1“ wäre die beste Position).

Eshel & Klein (1981) untersuchten in dieser Querschnittstudie 2199 Kinder, leider ohne Einbezug des Kindergartens. Sie liessen die Kinder sich selber "Noten" geben (Skala 1 bis 5). Für Mathematik ergaben sich Mittelwerte von 4.75 (Unterschichtkinder), respektive 4.76 (Mittelschicht) im ersten Schuljahr, die dann auf 4.05 und 4.38 (4.Klasse) zurückgingen. Dieser Rückgang ist relativ klein, gemessen an den immer noch sehr hohen Durchschnittswerten. Im Vergleich der ersten mit der zweiten Klasse ist der Rückgang für die Mittelschicht praktisch null, für die Unterschicht nicht signifikant.

Die Befunde von Stipek und Tannatt (1984) zeigen einen Abfall der Einschätzhöhe, der aber nicht signifikant ist ($p > 0.10$). Es sind jeweils zwei Jahrgänge zusammengefasst, das Durchschnittsalter der Gruppen liegt jeweils zwei Jahre auseinander.

Bei Beneson und Dweck (1986) ist dagegen ein leichter Anstieg vom Kindergarten zur ersten Klasse zu finden, anschliessend ein Abfall. Dies ergibt einen diskontinuierlichen Verlauf wie bei Helmke (1991), der Abfall vom Kindergarten zur vierten Klasse ist signifikant ($p < 0.01$).

Die Befunde von Frey und Ruble (1987) zeigen keinen signifikanten Abfall der Einschätzhöhe: die Differenz zwischen Lehrerurteil und Selbstbeurteilung ist gemessen an den Möglichkeiten einer Vierzigerskala generell gering und der Abfall im Bereich eines Zufallsergebnisses.

Bei Stipek und Daniels (1988) ist der Abfall vom Kindergarten zur vierten Klasse bei den Kindern, die Schulklassen mit häufigem Leistungsfeedback durch die Lehrkraft besuchen, kaum vorhanden: Die Kinder haben im Durchschnitt in der vierten Klasse noch die fast gleich hohe Einschätzung gemacht, wie im Kindergarten. Ein Abfall findet sich bei Kindern die wenig Leistungsfeedback erhalten. Insgesamt ist der Abfall der beiden Gruppen signifikant ($p < .001$) (vgl. auch: Stipek & Mc Iver, 1989).

Fazit: Bei der Einschätzung der Höhe der eigenen Leistung im Lesen zeigt sich in den beiden dargestellten Studien, dass die Höhe der Einschätzung vom Kindergarten zur ersten

Klasse abnimmt und in der zweiten Klasse noch fast gleich hoch ist wie ein Jahr zuvor. Die Leseerfahrung im ersten Schuljahr wirkt sich in diesen Studien interessanterweise negativ auf die Einschätzhöhe aus und sinkt bis in die mittleren Schuljahre nur noch unbedeutend. Was die allgemeine Schulleistungsfähigkeit betrifft ist die Befundlage unterschiedlich: Bei Stipek und Tannatt (1984) nimmt die Einschätzhöhe nicht signifikant ab, bei Beneson und Dweck (1986) ist sie diskontinuierlich, nimmt aber nur im Vergleich Kindergarten zu vierter Klasse signifikant ab und bei Stipek und Daniels (1986) wird die allgemeine Leistungskompetenz in der vierten Klasse noch fast gleich hoch eingeschätzt wie im Kindergarten. Die Abnahme der Höhe der Einschätzung bei Eshel und Klein (1981) im Fachbereich Mathematik ist ebenfalls nicht signifikant. Leider umfasst diese Studie aber den Kindergarten nicht. Die Genauigkeit der Einschätzung nimmt in den obigen Untersuchungen zu. Ab zweiter Klasse erreichen die Korrelationen signifikante Werte, während vorher fast kein oder nur ein sehr geringer Zusammenhang festgestellt werden kann. Nun zu den Befunden aus den neunziger Jahren (Tab.2-4-2):

Tabelle 2-4-2: Mittelwerte und Korrelationen von Leistungseinschätzungen (1990-1997)

Autor, Stichprobe		KG	1.KI	2.KI	3.KI.	4.KI	5.KI	6.KI	Beurteilungsbereich Normbezug
Helmke, 1991, N=100, pro Jahr	M r	2.2 0.15	3.1 0.16	1.9 0.32	2.3 0.34				Mathematik Soziale Norm, Skala: - 5 bis +5 „Ich bin am besten“ =5
Eccles et al., 1993 N=865, total	M M M		5.69 5.74 5.30	5.49 5.59 5.26		5.38 5.68 4.70			Mathematik Sport,allgemein Bodenturnen Normmix, Skala 1-7 „Ich bin sehr gut/ am besten“ = 7
Weinert, Stefanek, 1997 N=200 pro Jahr	M	2.96	3.12	2.80	2.75	2.52			Allgem. Leistungs- selbstbild , Soziale Norm, Skala: -5 bis +5 „ich bin am besten“ =5
Wigfield et al., 1997 N=200-300, pro Jahr	M		5.7 0.01	5.5 0.14	5.2 0.27	5.3 0.37	5.3 0.34	5.2 0.50	Mathematik Normmix, Skala: 1-7 „ich bin am besten“ =7

Die Befunde der Untersuchungen zeigen einen deutlichen Abfall der Höhe der Einschätzung von der ersten zur zweiten Klasse. Vergleicht man allerdings den Kindergarten mit der zweiten Klasse ist der Abfall sehr gering und nicht signifikant. Wieder findet sich diskontinuierliches: Die Untersuchungen fanden einen Anstieg der Selbsteinschätzung vom Kindergarten zur ersten Klasse mit anschliessendem Abfall.

Bei den Studien von Eccles et al. (1993) und von Wigfield et al. (1997) wurde die Einschätzung durch die Kinder über mehrere Bezugsnormen erhoben und daraus ein Mischwert errechnet. In diesen Untersuchungen sinken die Einschätzhöhen weniger stark als in den anderen Studien. Eine Mischung aus absoluter Leistungsbeurteilung und sozialem Vergleich kommt der Schulrealität ab zweitem Schuljahr meines Erachtens näher, während sich die jüngeren Kinder aufgrund ihrer kognitiven Entwicklung vermutlich vorwiegend am eigenen Lernfortschritt einschätzen.

Die Genauigkeit der Einschätzung nimmt auch in diesen Befunden zu, Helmke (1991) findet bei den Zweitklässlern eine relativ hohe, signifikante Korrelation.

Insgesamt ergibt sich aus diesen Befunden, dass mit einem signifikanten Abfall vom Kindergarten bis zur zweiten Klasse eher nicht zu rechnen ist und dass sogar ein diskontinuierlicher Anstieg vom Kindergarten zur ersten Klasse zu erwarten ist. Signifikante Korrelationen können vermutlich schon in der ersten, spätestens in der zweiten Klasse erwartet werden.

Im folgenden werden zuerst zwei Untersuchungen zitiert, deren Ergebnisse für die Fragestellungen der vorliegenden Arbeit nicht mehr genutzt werden konnten, deren Befunde bei der Datenerhebung im Kindergarten aber schon vorlagen. Daran schliesst eine weitere Untersuchung an, die nach der Datenaufnahme publiziert wurde. Dazu kommen zwei Untersuchungen aus einem Lizentiatspool der Universität Bern. Im Rahmen einer grossen Querschnittuntersuchung wurden im November 2001 Kinder vom zweiten bis zum achten Schuljahr bezüglich ihrer Selbsteinschätzung in diversen schulischen Leistungsbereichen untersucht. Die Datenerhebung erfolgte vier Monate nach der letzten Datenaufnahme für die vorliegende Arbeit. Als Instrument zur Erfassung der Selbsteinschätzung wurde eine etwas veränderte Form der Treppenstufenskala aus der vorliegenden Untersuchung eingesetzt. Die Vereinfachung wurde nötig, weil die Einschätzungen nicht im Einzelinterview erfolgten, sondern im Rahmen einer klassenweisen Instruktion. Dadurch entfiel insbesondere die Möglichkeit, die Einschätzungen eines Kindes durch Rückfragen zu verifizieren.

Tabelle 2-4-3: Mittelwerte und Korrelationen von Leistungseinschätzungen (1998-2002)

Autor, Stichprobe		Vor KG	KG	1.KI	2.KI	3.KI.	4.KI	5.KI	6.KI	Beurteilungsbereich Normbezug
Measelle et al., 1998 N= 73 – 97 pro KI	M R	0.49	0.48 0.19	0.44 0.24						Allgemeine Schulfähigkeit Im Durchschnitt Normmix, Skala: -1 bis +1 „Ich bin sehr gut“ = +1
Marsh, Craven, Debus, 1998 N= 127–137 pro KI.	M		4.39 4.18 4.17	4.37 4.18 4.12	4.39 4.14 4.02	4.27 4.09 3.91				Sportlichkeit Mathematik Allg. Schulleistung SDQ –Norm, Skala 1-8 „Ich bin sehr gut“ = 8
Valeski & Stipek, 2001, N=KG.:225, 1.KI.:147	M R		4.04 0.10	4.39 0.34						Mathematik Normmix, Skala: 1-5 „Ich bin am besten“ =5
Demény, 2002 N= 106 –125	M R				10.0 0.27	9.7 0.32	8.6 0.16	7.9 0.42	8.4 0.54	Mathematik Soziale Norm, Skala: 1-15 „Ich bin der beste“ = 15
Ingold, 2002 N=106-125	M R				11.7 0.36	9.9 0.47	8.1 0.67	8.5 0.59	8.5 0.52	Sport Soziale Norm, Skala:1-15 „Ich bin die beste“=15

Das Erhebungsverfahren von Measelle et al. (1998) basiert auf dem Berkeley Puppet Interview (BPI), bei dem die Kinder die Fragen von Puppen gestellt erhalten und selber auch durch Puppen antworten. Dieses sehr kindernahe Verfahren war notwendig, weil es sich um sehr junge Kinder handelte: Vorschulalter: 4.6 Jahre, Kindergarten: 5.9 Jahre, erste Klasse: 6.9 Jahre. In dieser Untersuchung zeigte sich, dass die Kinder schon sehr gut zwischen verschiedenen Bereichen des Selbstkonzeptes unterscheiden können. Die Abnahme der durchschnittlichen Einschätzhöhe ist nicht signifikant, die Genauigkeit erreicht schon in der ersten Klasse ein signifikantes Niveau.

Sehr erstaunlich sind die Ergebnisse der umfangreichen Untersuchung von Marsh, Craven, Debus (1998), in der eine stark gekürzte Version des SDQ (Selbstbeschreibungsbogen) zum Einsatz kam. Das Absinken der Einschätzhöhe ist im Sport und der Mathematik nicht signifikant, in der Beurteilung der allgemeinen Schulleistungsfähigkeit hingegen schon ($p=0.04$). Interessant an dieser Untersuchung sind zudem zwei weitere Befunde: Bereits sehr früh unterscheiden die Kinder differenziert zwischen verschiedenen Bereichen des Selbstkonzeptes und die Einschätzung ist intraindividuell bereits sehr früh recht stabil.

Bei Valeski & Stipek (2001) ist der Anstieg der Höhe der Einschätzung signifikant, die Korrelation ist in der ersten Klasse signifikant und erreicht im Vergleich mit den anderen Untersuchungen in Mathematik den höchsten Wert.

Demény (2002) setzte für die Leistungserhebung in Mathematik ebenfalls Aufgaben aus dem KAB-C ein, allerdings wurde für das zweite Schuljahr aus zeitökonomischen Gründen nur eine verkürzte Form benutzt. Die Einschätzungen weichen, mit Ausnahme des fünften Schuljahres, signifikant vom erwarteten Mittelwert (8) ab. Die Einschätzhöhe nimmt über die Schuljahre hin signifikant ab. Mit Ausnahme des vierten Schuljahres sind alle Korrelationen signifikant. Die Zunahme der Schätzgenauigkeit ist nicht signifikant.

Ingold (2002) operationalisierte die Leistungen im Sport durch drei Aufgaben (Stosskraft mit Arm, Ausdauer und Weitsprung). Vom zweiten bis zum vierten Schuljahr weichen die Einschätzhöhen signifikant von einem mittleren Wert (8) ab, in den weiteren Schuljahren sind die Abweichungen nicht mehr signifikant. Die Korrelationen sind signifikant, die Genauigkeit der Selbsteinschätzungen nimmt über die Klassenstufen signifikant zu.

Die ersten beiden Studien zeigen, dass die Einschätzhöhe vom Kindergarten zur zweiten Klasse konstant gleich überdurchschnittlich hoch bleibt und zwar im Sport und in Mathematik. Nur in der allgemeinen Schulleistungsfähigkeit ergibt sich eine signifikante Abnahme. Die Schätzgenauigkeit ist niedrig und erreicht im ersten Schuljahr knapp einen signifikanten Wert. Die beiden Untersuchungen von Ingold (2002) und Demény (2002) beginnen erst mit der Schulstufe, die in der vorliegenden Arbeit den Endpunkt der Messung markiert, dem zweiten Schuljahr. Zu diesem Zeitpunkt schätzen sich die Kinder signifikant überdurchschnittlich hoch ein. Vom zweiten Schuljahr an nimmt die Einschätzhöhe bis zum fünften Schuljahr sowohl im Sport wie in Mathematik signifikant ab. Die Schätzgenauigkeit ist recht hoch und schon im zweiten Schuljahr signifikant, sie nimmt aber nur im Sport signifikant zu. Weil vom fünften Schuljahr an auf den Übertritt in eine höhere Schule hin gearbeitet wird, könnte die resultierende persönliche Verunsicherung der Kinder erklären, wieso die Schätzgenauigkeit in der Mittelstufe in Mathematik so tief ist.

Insgesamt ergibt sich aus den Befunden, dass die Höhe der allgemeinen Leistungseinschätzung generell zwar sinkend ist, allerdings im Vergleich von Kindergarten mit der zweiten Klasse im Leistungsbereich Mathematik praktisch gleich hoch bleibt. Die älteren Befunde beruhen mit einer Ausnahme (Eshel & Klein, 1981) auf kleinen Stichproben pro gemessenem Jahrgang. Die gefundenen Verläufe der Höhe der Selbsteinschätzung konnten deshalb nur tendenziell bestätigt werden, die gemessenen Abnahmen erreichten kein Signifikanzniveau. Bei Eshel und Klein war die Stichprobengröße ausreichend, es fehlten dort aber Befunde zum Kindergarten, was insbesondere schade ist, weil sich in den anderen Studien im Über-

gang vom Kindergarten zur ersten Klasse mit einer Ausnahme steigende Werte fanden. Die neueren Befunde stammen aus deutlich grösseren Stichproben. Bei den Untersuchungen von Eccles et al., Wigfield et al. und Valeski und Stipek wurde die Genauigkeit der Einschätzung durch mehrere Normvergleiche eingeschätzt, nämlich an der absoluten eigenen Leistungsfähigkeit („Das kann ich gut“), im Vergleich mit anderen Fächern („Mathematik kann ich besser als Sport“) und an der sozialen Norm („Ich kann das schlechter als andere“). Aus diesen Antworten wurde dann die Einschätzhöhe in einzelnen Fächern berechnet (Normmix). Demgegenüber machen die deutschen Studien einen sozialnormlastigeren Vergleich: „Wie gut bist du *im Vergleich zu den anderen Schülern?* („Sehr gut“, „gut“, „manchmal gut und manchmal weniger gut, ...“). Auch hier wird aber nach der absoluten Leistungsqualität gefragt, es handelt sich also nicht um eine reine Rangreihe („*Wie gut* bist du, gemessen an den Mitschülerinnen und -schülern?“). Über alle drei analysierten Gruppen von Befunden gesehen, kann darauf geschlossen werden, dass ein leichter Anstieg der Einschätzhöhe vom Kindergarten zur ersten Klasse erfolgt und die Einschätzhöhe dann bis zur zweiten Klasse leicht sinkt. Ob sich Mathematik und Sport über die Jahre in der Einschätzhöhe im Schuleingangsalter unterscheiden, kann aufgrund der analysierten Befunde nicht schlüssig beantwortet werden. Die Genauigkeit der Einschätzung, gemessen als Korrelation von Leistung und Einschätzung, dürfte in der zweiten Klasse erstmals signifikant werden, vielleicht schon in der ersten. Einen substanziellen Wert (höher als etwa 0.4) erreicht sie aber erst in den mittleren Jahren der Grundschule. Insgesamt kann mit einer Zunahme der Genauigkeit gerechnet werden, wenn diese auch nicht immer signifikant ist. Die neueren Untersuchungen kommen zum gleichen Ergebnis, im Fachbereich Mathematik nimmt die Einschätzhöhe:

- von der ersten zur zweiten Klasse deutlich ab
- vom Kindergarten zur ersten Klasse zu
- im Vergleich von Kindergarten mit zweiter Klasse wenig ab

Eccles et al. (1993) bestätigen zudem den wichtigen Befund von Marsh et al. (1984), dass Kinder schon sehr früh, nämlich im Kindergarten, zwischen verschiedenen Leistungsbereichen unterscheiden können, wenn es darum geht, zu sagen, was man gut kann und was weniger. Dies ist insofern bedeutsam, als die Überschätzungstendenzen bei fachspezifischen Beurteilungen in den zitierten Studien deutlich tiefer waren, als bei der Beurteilung der generellen schulischen Leistungsfähigkeit. Die Genauigkeit der Selbsteinschätzung nahm demgegenüber konstant zu, blieb aber bis in die zweite Klasse tief. Substanziell stieg sie in der dritten Klasse deutlich an.

Das Leistungsselbstbild in Mathematik ist vermutlich am Ende der zweiten Klasse nur wenig tiefer, als im Kindergarten. Die Kinder haben von sich selbst eine konstant zu hohe Leistungsmeinung. Auf die Leistungsfähigkeit und das Lernselbstvertrauen wirkt sich das positiv aus. Zumindest von dieser Seite wäre deshalb eine positive Wirkung auf die Höhe der Lernfreude zu erwarten. Neben dem Vertrauen in die eigene Leistungskraft ist das Lernenwollen für den Lernerfolg auch wichtig, der Lernerfolg wiederum beeinflusst das Leistungsselbstbild. Wie wir bereits gesehen haben, ist die Lernfreude während der Schuleingangsphase bei allen Kindern hoch und nur für die Mädchen leicht sinkend (aber immer noch überdurchschnittlich).

2.5 Unterschiedliche Einschätzhöhe bei Knaben und Mädchen

Die geschlechtergetrennte Betrachtungsweise wurde bisher nicht aufgenommen, nicht zuletzt weil die geringen Stichprobengrößen bei vielen Untersuchungen eine solche nicht erlaubten, dann aber auch, weil die bisher zitierten Befunde für die Schuleingangszeit kaum signifikante Unterschiede belegen. Für die späteren Schuljahre sind aber Geschlechtsunterschiede sowohl bezüglich Selbsteinschätzung gut belegt (Eccles, 1984; Wigfield et al., 1991): Sie zeigen, dass die Unterschiede zwischen den Geschlechtern im Schulalter ab etwa 5. Klasse generell zunehmen: Mädchen schätzen sich in Musik und Sprache stärker ein, Knaben in Mathematik und Sport. Das entspricht teilweise auch dem „common sense“ der Gesellschaft, insbesondere beim Sport. Dort treten Frauen und Männer als Erwachsene in geschlechtsgetrennten Gruppen gegeneinander an. Diese Unterscheidung lässt sich im Sport auch aus den gemessenen Leistungsunterschieden begründen.

Heikler ist die Diskussion darüber, ob in Mathematik und Sprache nun geschlechtsspezifische Leistungsunterschiede sozial konstruiert sind oder nicht. Marsh et al. (1984) befragten pro Schuljahr zwischen 103 bis 251 Kinder zum Selbstkonzept, von der zweiten bis zur fünften Klasse. Bereiche waren unter anderem „physikalische Fähigkeiten“ (Sport), Lesen und Mathematik. Knaben schätzten sich im Sport bereits in der zweiten Klasse signifikant leistungsstärker ein als Mädchen. Mädchen schätzten sich umgekehrt im Lesen signifikant stärker ein als Knaben. Bezüglich der gesamtschulischen Leistungsfähigkeit schätzten sich Mädchen vom zweiten Schuljahr an konstant höher ein als Jungen (wenn auch nur knapp signifikant). In Mathematik waren keine signifikanten Unterschiede messbar. Die oben berichtete Zunahme der Geschlechtsunterschiede in der Einschätzung der Leistungen fand sich in der Untersuchung von Marsh et al. nicht, vom zweiten zum fünften Schuljahr blieb der Unterschied zwischen den Geschlechtern stabil.

Befunde zum ersten Schuljahr sind spärlich. Eccles (1993) fand allerdings auch in Mathematik signifikante Unterschiede in der Einschätzung, die Stichprobe umfasste Kinder vom ersten bis vierten Schuljahr. Im Sport fand Eccles (1993) Unterschiedliches: In Bodengymnastik schätzen sich Mädchen höher ein, im Ballwerfen und -fangen hingegen die Knaben. Die Unterschiede in den Befunden zur Mathematik sind allerdings klein und auf die Frage, welche Fächer die Kinder der Reihe nach als ihre besten Leistungsbereiche bezeichnen, sagen:

Knaben: Sport, Lesen, Mathematik, Instrument spielen

Mädchen: Lesen, Sport, Mathematik, Instrument spielen

Für Knaben und Mädchen erscheint Mathematik erst an dritter Stelle, beide glauben, dass sie sowohl im Sport wie im Lesen besser sind als in Mathematik. Für Knaben ist die Mathematik in ihrer Wahrnehmung nicht leichter als Sport oder Lesen. Die Befunde von Eccles unterscheiden sich noch in einem anderen Punkt, die Geschlechtsunterschiede sind in allen gemessenen Bereichen in der ersten Klasse kleiner als in der vierten. Ob die Unterschiede in Mathematik allerdings bereits schon im ersten Schuljahr signifikant sind, bleibt offen.

Wigfield et al. (1997) fanden ebenfalls signifikante Unterschiede in der Selbsteinschätzung in Mathematik, diese Stichprobe enthält Kinder aller Jahrgänge von der ersten bis zur sechsten Klasse. Allerdings beziehen sich diese Unterschiede auf die Gesamtpopulation, also auf den erwarteten Durchschnitt für die Grundschuljahre. In der Studie von Valeski und Stipek (2001), die nur Kindergarten und erstes Schuljahr umfasst, fanden sich für die beiden Populationen keine Geschlechtsunterschiede bezüglich Leistungseinschätzung in Mathematik. Valeski und Stipek benützten ein Messinstrument, das eine Weiterentwicklung desjenigen von Wigfield et al. war, abgestimmt auf die Möglichkeiten der Kindergartenkinder.

Tiedemann und Faber (1995, 1994) untersuchten, ob die beobachteten Unterschiede der Selbst(über-)schätzung der Mathematikleistungen auf das Leistungsfeedback der Lehrperson zurückgehen könnte. Von der ersten bis zur vierten Klasse erhielten die Mädchen allerdings durch die Lehrpersonen bessere Beurteilungen, obwohl ihre Leistungen dies nur im ersten Schuljahr gerechtfertigt hätten. Die Leistungen der Mädchen in Mathematik nähern sich in den folgenden Jahren denjenigen der Knaben an, im vierten Schuljahr sind sie vergleichbar. Vom Vorschulalter bis zur zweiten Klasse unterscheiden sich die Leistungen nur in der ersten Klasse signifikant und zwar zugunsten der Mädchen.

Mädchen werden von den Lehrkräften also besser beurteilt und sind leistungsmässig leicht besser, ihre tiefere Leistungsselbstwahrnehmung kann also nicht auf dem Einfluss geschlechtstypisierender Leistungsbeurteilung durch die Lehrpersonen oder auf erlebter schwächerer Leistung beruhen. Mädchen schätzen sich zu einem Zeitpunkt schwächer ein, in dem ihre faktischen Leistungen denen der Knaben ebenbürtig sind. Bereits in der

3./4.Klasse tendieren Mädchen stärker zu ungünstigen Kausalattributionen, wenn es darum geht Lernerfolg, respektive Misserfolg in Mathematik zu erklären (Tiedemann & Faber, 1995). Lehrkräfte ihrerseits haben insgesamt die Erwartung, dass Mädchen in diesem Alter weniger ausgeprägte Problemlösefähigkeiten haben, dafür fleissiger sind in Mathematik, ihr Leistungspotential also eher ausschöpfen, als Knaben (Tiedemann, 1995). Solche geschlechtsspezifischen Kompetenzerwartungen der Lehrkräfte sind vermutlich auch bereits im Schuleingangsalter vorhanden. Ob sie die Selbstbeurteilung allerdings beeinflussen können, ist eher wenig wahrscheinlich, denn die Kinder orientieren sich bei der Beurteilung der eigenen Leistungsfähigkeit in diesem Alter ja am Leistungsfortschritt.

Cole et al. (1998) untersuchten an 807 Kindern, ob sich bezüglich der schulischen Leistungseinschätzung Geschlechtsdifferenzen ergeben. In dieser Untersuchung überschätzten sich die Knaben durchgängig von der dritten bis zur sechsten Klasse, während sich die Mädchen unterschätzen. Die Geschlechtsdifferenz war in der vierten Klasse erstmals signifikant ($p < 0.05$), ab dem sechsten Schuljahr noch ausgeprägter ($p < 0.01$).

Demény (2002) fand in Mathematik von der zweiten bis zum sechsten Schuljahr keine signifikanten Leistungsunterschiede zwischen Mädchen und Jungen, hingegen kommt es zu signifikanten Unterschieden in der Einschätzung der Leistung: Vom zweiten zum vierten Schuljahr schätzten sich die Knaben signifikant höher ein, nachher wird der Unterschied kleiner, bleibt aber signifikant. Bei Ingold (2002) waren die Geschlechtsunterschiede in der Einschätzung ebenfalls signifikant ($p < 0.01$). Im Sport waren die Leistungen der Knaben allerdings in allen Aufgaben ebenfalls signifikant höher. Insgesamt waren es die Mädchen, die sich genauer einschätzten.

Lesen, Schreiben und Mathematik sind zentrale Lernfelder der Schule, bei denen die Gesellschaft davon ausgeht, dass es zwischen den Geschlechtern Unterschiede in der Leistungsfähigkeit gibt. Wie die Befunde von Tiedemann und Faber zeigen, neigen Mädchen zu einem negativen Bild eigener Leistungsfähigkeit in Mathematik, bevor ihre Leistungen effektiv hinter denen der Knaben zurückzuliegen beginnen. Die Befunde von Wigfield et al. und Stipek zeigen, dass Kinder sehr früh in der Lage sind, einzelne Leistungsbereiche klar zu unterscheiden und sich auch entsprechend unterschiedlich einzuschätzen. Im Sport konnten Kinder zwischen allgemeiner sportlicher Leistungsfähigkeit und ihren Leistungen im Ballwerfen oder in Bodengymnastik bereits in der ersten Klasse unterscheiden: Mädchen schätzen sich dabei in Bodengymnastik besser ein, Knaben dafür im Ballwerfen und der allgemeinen sportlichen Leistungsfähigkeit. Die Leistungsfähigkeit im Sportbereich zu messen und mit den gemachten Einschätzungen zu überprüfen wäre deshalb interessant, Die Gesellschaft er-

wartet wahrscheinlich in keinem anderen Leistungsbereich so geschlechtsspezifische Unterschiede wie im Sport. In keiner Sportart treten die Geschlechter gegeneinander im Wettbewerb an, nicht einmal im Schach oder im Schiesssport.

Vermutlich ist der Sport derjenige schulische Lernbereich, indem die Leistungsfähigkeit in den verschiedensten Sparten am besten untersucht wurde. Die Ergebnisse in diversen Leistungsbereichen zeigen dabei, dass im Alter zwischen sechs und neun Jahren wenig Leistungsunterschiede festgestellt werden können. Gaschler (1998) findet Unterschiede im Vorschulalter (5-6 Jahre) in der Gesamtkörperkoordination, also beispielsweise beim Balancieren und im Hüpfen. Diese Unterschiede fanden auch Kiphard und Schilling (1974). Mädchen waren in diesem Alter leicht überlegen, allerdings nur im Balancieren rückwärts signifikant. Im achten Lebensjahr werden dann die Knaben leicht besser, aber auch nicht signifikant. Im Hüpfen, einer Aufgabe zur Überprüfung der Kraft, sind die Mädchen bis in die erste Klasse signifikant überlegen, in der zweiten Klasse haben die Knaben leicht überholt (Kiphard & Schilling, 1974). Bezüglich Schnellkraft und Ausdauer sind keine Leistungsunterschiede im Schuleingangsalter zu erwarten (Althaus & Bühler, 1999). Es ergibt sich also im Kindergarten eine leichte Überlegenheit der Mädchen, die sich bis in die zweite Klasse verliert. Thomas und French (1985) untersuchten Geschlechtsunterschiede in einer Metaanalyse von 64 Untersuchungen mit einem insgesamt von 702 verschiedenen Leistungsitems im Sportbereich. Die Altersspannweite der Metaanalyse reicht vom dritten zum 19. Lebensjahr. Motorische Leistungsfähigkeit wurde in 20 Teilbereichen analysiert: Nur in fünf dieser Bereiche waren vor der Pubertät signifikante Leistungsunterschiede feststellbar. Am Ende der zweiten Klasse finden sich deutliche Überlegenheiten der Knaben nur im Ballwerfen. Kaum Unterschiede zwischen Kindergarten und zweiter Klasse finden sich bei Bewegungskoordinationaufgaben, sowohl fein- wie grobmotorisch, in der Balancierfähigkeit (minimale Überlegenheit der Mädchen) oder im Hochsprung. Thomas und French schliessen daraus, dass die unterschiedliche Erwartung von Knaben und Mädchen an die eigene Leistungsfähigkeit vermutlich durch Gleichaltrige und Eltern vermittelt werde. Wenn Mädchen gesellschaftliche Erwartungen (vermittelt durch Eltern und Gleichaltrige) im Sinne von schlechterer Leistungsfähigkeit bereits früh zu übernehmen beginnen, wie es Tiedemanns Befunde nahelegen, dann müsste das im Bereich Sport ausgeprägter der Fall sein als in Mathematik: Schon im Kindergartenalter sind die Kinder in der Lage zu begründen, wieso es keine geschlechtsgemischten Wettkämpfe gibt. Diesbezügliche Unterschiede in der Leistungsfähigkeit in den Bereichen Mathematik und Sprache hingegen werden im gesellschaftlichen Alltag kontrovers diskutiert und, weil sie sehr relevant für die Bildungschancen sind, eher heruntergespielt.

Im Alltag von Kindergarten und Schule ist der Sport das einzige „Fach“, welches sowohl im Kindergarten und der Schule unter vergleichbaren Bedingungen erworben wird: Der Sportunterricht findet meisten schon im Kindergarten in der Schulturnhalle statt und das Sportlehrmittel beginnt bereits mit dem Kindergartenalter. Im Leistungsbereich Sport ergibt sich also vom Kindergarten zur zweiten Klasse eine kontinuierliche Entwicklung. In Mathematik und Sprache gilt das nicht. Hier setzen gezielter Fachunterricht und die Arbeit mit einem Lehrmittel erst in der ersten Klasse ein. Gefördert wird das Rechnen aber bereits im Kindergarten, wenn auch eher beiläufig, so doch im Hinblick auf die rechtzeitige Erkennung allfälliger Entwicklungsrückstände. Für die Untersuchung von Geschlechtsunterschieden ist deshalb die Auswahl der beiden Fächer Sport und Mathematik günstig, denn Leistungsunterschiede sind im Schuleingangsalter sehr klein und vermutlich sind es die Mädchen, die eher besser sind. Es sind aber auch beides Fächer, bei denen häufig angenommen wird, Jungen würden die besseren Leistungen erbringen, die Erwartungen liegen deshalb anders als die tatsächlichen Leistungserfahrungen. Das heisst, die tieferen Leistungseinschätzungen der Mädchen sind sozial konstruiert. Ob sich bereits im Schuleingangsalter Geschlechtsunterschiede finden lassen, wird sich weisen. Die bisherigen Befunde sprechen dafür, dass sich Knaben höher einschätzen als Mädchen.

2.6. Fragestellungen und Hypothesen

2.6.1 Verlaufsgestalt der Selbsteinschätzung

Die zum Zeitpunkt der Vorbereitung der Erhebung erster Daten für die vorliegende Untersuchung bereits publizierten Befunde zeigen, dass die Kinder im Kindergarten- und Grundschulalter im Durchschnitt ihre Leistung überdurchschnittlich hoch einschätzen. In einer 6-Cluster-Lösung fand Helmke (1991) zu Beginn der zweiten Klasse bei keinem einzigen Cluster eine Tendenz zur unterdurchschnittlichen Einschätzung der eigenen Leistung in Mathematik (gemessen an der sozialen Bezugsnorm). Auch am Ende der zweiten Klasse, also ein halbes Jahr später, waren es erst 73 von 757 Kindern, die zu einer leicht unterdurchschnittlichen Einschätzung neigten ($z = -0.1$). In den Studien wurde die Höhe der Einschätzung, wie bereits diskutiert, an verschiedenen Massstäben gemessen, wie die Tabellen 2-4-1 bis 2-4-3 zeigen. Manchmal schätzen sich die Kinder komparativ ein, sie beurteilen ihre Leistung im Vergleich zu den anderen Schülerinnen und Schülern ihrer Klasse. Bei Helmke (1992, S.210) wurde beispielsweise gefragt: „Wie gut bist du in Mathematik im Vergleich zu

den anderen deiner Klasse?“ Das Kind konnte auswählen zwischen „am schlechtesten“, „deutlich unter dem Durchschnitt“, usw. In anderen Untersuchungen orientierten sich die Kinder statt an der sozialen, an der absoluten Norm: „Ich kann sehr gut rechnen“ (Eshel & Klein, 1981). Neuere Untersuchungen erheben mit verschiedenen Skalen beide Normbezüge (vgl. Tab.2-4-2 und 2-4-3, dort als „Normmix“ bezeichnet).

Grundsätzlich ist davon auszugehen, dass sich die Kinder zu hoch einschätzen, ob sie sich nun am sozialen oder am absoluten Massstab messen, im Schuleingangsalter findet aber eine Verschiebung der Präferenz statt, die Kinder orientieren sich zunehmend am sozialen Massstab. Eine Messskala, die beide Bezugsnormen kombiniert ist deshalb günstig (Hypothese 1.1). Zusätzlich sollen die Kinder ihre Leistungsfähigkeit, dort wo dies möglich ist, auch ohne den Vergleich mit der Klasse vornehmen können. Das kann im Sport der Fall sein, wenn es beispielsweise gilt, einzuschätzen, wie weit oder wie hoch man springen kann (Hypothese 1.2). In Mathematik ist eine solche Einschätzung im Schuleingangsalter dagegen kaum möglich, absolute Leistungsstandards werden durch die Lehrkräfte erst später ausreichend transparent und häufig kommuniziert (vgl. Kap. 2.2).

Hypothese 1: Kinder schätzen ihre Leistungsfähigkeit im Durchschnitt im Kindergarten, der erste und zweiten Klasse zu hoch ein.

1.1.: Kinder aller Klassenstufen schätzen sich selbst in Mathematik und Sport höher ein, als sie den Durchschnitt ihrer Klasse einschätzen.

1.2.: Im Durchschnitt schätzen die Kinder die eigene Leistungsfähigkeit im Sport in jedem Jahr, gemessen an der tatsächlichen Leistung, zu hoch ein.

Die Längsschnittstudien der letzten 10 Jahre zeigen für den Fachbereich Mathematik eine Zunahme der absoluten Einschätzhöhe der eigenen Fähigkeiten vom Kindergarten zur ersten Klasse und eine anschließende Abnahme. Bei Helmke (1991) ist in der zweiten Klasse die Einschätzhöhe nicht signifikant unterhalb derjenigen im Kindergarten und in der dritten Klasse sogar wieder höher als im Kindergarten. Hingegen findet er einen signifikanten Anstieg der durchschnittlichen Einschätzhöhe vom Kindergarten zur ersten Klasse und einen anschließenden signifikanten Abfall der Einschätzhöhe bis zur zweiten Klasse. Eine lineare Entwicklung ergibt sich in seiner Untersuchung über diesen Zeitraum nicht. Trotzdem vermuten wir, dass die Einschätzhöhe in unserer Untersuchung zunimmt, weil die Kinder ja erst im ersten Schuljahr Mathematikunterricht erfahren (dies im Gegensatz zu den Kindern in den

amerikanischen Studien, die jeweils schon im Kindergarten Mathematikunterricht haben). Da sie sich bei der Beurteilung vorwiegend am eigenen Leistungsfortschritt messen, müssten sie sich in der ersten Klasse höher einschätzen: Sicher haben sie im ersten Schuljahr mehr gelernt, als im Kindergarten. In der zweiten Klasse ist ebenfalls von einer höheren Einschätzung als im Kindergarten auszugehen, denn der Zuwachs an Leistung ist vermutlich ebenfalls grösser als im Kindergarten. Allerdings erwarten wir auch, dass nun, neben der intraindividuellen, die interindividuelle Bezugsnorm bei der Beurteilung beigezogen wird. Das müsste zu einer durchschnittlich weniger hohen Einschätzung der eigenen Leistung führen, als in der ersten Klasse (sofern die Leistungszunahme im ersten und zweiten Schuljahr gleich sind).

Im Sport verfügen wir über keine Längsschnittdaten zum Verlauf der Einschätzhöhe, die bereits mit dem Kindergarten beginnen würden. Von der ersten zur zweiten Klasse nimmt die Einschätzhöhe leicht (nicht signifikant) ab, wie Wigfield et al. (1997) berichten. Ihre Befunde zu Mathematik und Sport (erste bis sechste Klasse, im Jahresabstand gemessen) ergeben im Vergleich des ersten mit dem vierten Schuljahr für beide Fächer ungefähr die gleiche Abnahme der Überschätzung. Es kann nur vermutet werden, dass der Verlauf von der ersten zur zweiten Klasse ebenfalls eine sinkende Tendenz aufweist. Schwieriger ist der Rückschluss auf den Kindergarten. Das Kind der Schuleingangsstufe misst ja die eigene Leistung in Schulfächern vorwiegend intraindividuell, also am eigenen Fortschritt, erst gegen Ende der zweiten Klasse gewinnt die soziale Bezugsnorm an Bedeutung. Im Sport erleben Kinder schon früh, dass andere schneller rennen können, höher springen, Velofahren beherrschen, Skifahren können, usw. Der soziale Vergleich beginnt im Sport also vermutlich früher als in Mathematik. Zudem verfügt das Kind im Kindergarten im Sport über vergleichende Informationen aus der Gleichaltrigengruppe: Es hat regelmässig Sportunterricht mit anderen und sieht, beispielsweise im Staffettenlauf, schnell, dass andere schneller oder langsamer rennen können. Durch den regelmässigen Sportunterricht im Kindergarten kann der Lernfortschritt dann in der Schule auch nicht so stark zunehmen wie in Mathematik, wo der Unterricht ja erst in der Schule beginnt. Von der Leistungsfähigkeit her ist nach Gaschler (1998) im Alter von fünf Jahren ein grösserer Leistungszuwachs erfolgt, dann bleibt die Leistungsentwicklung bis ins elfte Lebensjahr gleich kontinuierlich zunehmend. Dies alles führt deshalb im Sport während der Schuleingangsphase vermutlich weder zu einem An- und Abstieg der Einschätzhöhe (wie er in Mathematik erwartet wird) noch zu einem signifikanten Anstieg der Einschätzhöhe vom Kindergarten im Vergleich zur zweiten Klasse. Erwartet wird allenfalls ein Trend zur Abnahme der Einschätzhöhe.

Hypothese 2: Die Entwicklung der Einschätzhöhe eigener Leistung verläuft nicht in allen Fachbereichen gleich.

2.1.: In Mathematik schätzen sich die Kinder in der ersten Klasse höher ein als im Kindergarten und als in der zweiten Klasse.

2.2.: Die Einschätzhöhe in Mathematik ist in der zweiten Klasse höher als im Kindergarten.

2.3.: Die Einschätzhöhe der eigenen Leistungsfähigkeit im Sport ist im Kindergarten gleich hoch wie in der ersten Klasse und in der zweiten Klasse (Nullhypothese).

2.4.: Die Einschätzhöhe der eigenen Leistungsfähigkeit im Sport ist im Kindergarten gleich hoch wie in der zweiten Klasse (Nullhypothese).

2.6.2 Die Schätzgenauigkeit nimmt zu

Hansford und Hattie (1982) fanden in ihrer Metaanalyse von über 120 Studien durchschnittliche Korrelationen von 0.12 für die Vorschule, 0.20 für die Grundschuljahre und 0.27 für die Oberstufe, wenn als Aussenkriterium Leistungstests verwendet wurden. Solch tiefe Werte deuten auf einen sehr schwachen Zusammenhang zwischen Selbstbeurteilung und Aussenkriterium. Allerdings lässt sich bereits ab der ersten Klasse eine Zunahme feststellen. In Bezug auf die Zunahme der jährlichen Schätzgenauigkeit (accuracy) vom Kindergarten zur zweiten Klasse finden sich keine Untersuchungen mit einem durchgehend gleichbleibenden Aussenkriterium. Besonders bei Helmke (1991) ist eine sprunghafte Zunahme der Schätzgenauigkeit in der zweiten Klasse festzustellen (auf $r=0.34$), die gerade mit einem solchen Wechsel des Aussenkriteriums einher geht. Die Befunde deuten darauf hin, dass die Schätzgenauigkeit in Mathematik vom Kindergarten zur ersten Klasse und zur zweiten Klasse hin ansteigt. Die Erklärung dafür ist plausibel: Erst mit dem Eintritt in die Schule beginnt der regelmässige Mathematikunterricht.

Ein Vergleich mit einem Fachgebiet, in dem das Kind schon Wettbewerbssituationen unter Gleichaltrigen kennt, könnte zeigen, inwiefern die Ungenauigkeit der Schätzung von der kognitiven Entwicklung abhängig ist und nicht von den fehlenden Erfahrungsmöglichkeiten. Das geeignete Fach dazu ist der Sport: In diesem Lernbereich hat das Kind seit frühester

Kindheit Erfahrung mit Erfolg und Misserfolg. Es weiss zudem eines: Ältere Menschen sind grundsätzlich leistungsstärker. Im Kindergarten erlebt es nun erstmalig Sportunterricht unter „fairen“ Bedingungen, das heisst unter Gleichaltrigen. Da der Lehrplan für Kindergarten und die ersten beiden Primarschuljahre im Kanton Bern die gleichen Ziele verfolgen und im gleichen Lehrmittel beschrieben sind, findet in diesem Lernbereich ein nahtloser Übergang in die Schule statt, einzig mit dem Unterschied, dass es in der Schule mehr Unterricht im Sport gibt als im Kindergarten. Im Kindergarten hat das Kind also bereits Erfahrung im unterrichtlichen Lernen von Sport, die Schätzgenauigkeit ist deshalb vermutlich auf dieser Stufe grösser als in Mathematik.

Hypothese 3: Die Schätzgenauigkeit zwischen Einschätzung und Leistung nimmt zu.

3.1.: Die Schätzgenauigkeit nimmt in Mathematik vom Kindergarten bis zur ersten, und von der ersten zur zweiten Klasse zu.

3.2.: Die Schätzgenauigkeit nimmt in Mathematik vom Kindergarten zur zweiten Klasse zu.

3.3.: Die Schätzgenauigkeit nimmt im Sport vom Kindergarten zur ersten, und von der ersten zur zweiten Klasse zu.

3.4.: Die Schätzgenauigkeit nimmt im Sport vom Kindergarten zur zweiten Klasse zu.

2.6.3 Geschlechtsunterschiede

Die bisherig zitierten Befunde haben klar auf Geschlechtsunterschiede hingedeutet, obwohl Hansford und Hattie (1982) beispielsweise in Bezug auf die Schätzgenauigkeit in ihren Meta-Analysen keine signifikanten Unterschiede fanden. Die meisten Untersuchungen beziehen sich allerdings auf das höhere Grundschulalter. Die Frage, inwiefern bereits in den ersten Schuljahren geschlechtsbezogene Unterschiede in der Höhe der Selbstbeurteilung zu finden sind, ist widersprüchlich zu beantworten. Marsh et al. (1984) fanden bereits bei Zweitklässlern Unterschiede: Mädchen überschätzten sich im Lesen, Knaben hingegen im Sport. Keine Unterschiede fanden sich in Mathematik. Bei Eccles et al. (1993) fanden sich dagegen höhere Einschätzungen bei Knaben im Sport und in Mathematik, allerdings wurden hier die Ergebnisse von erster, zweiter und vierter Klasse im gemeinsamen Durchschnitt

gemessen. Tiedemann (1995) konnte für dritte / vierte Klasse bereits deutliche Unterschiede finden: so haben Mädchen in seiner Untersuchung ein vermindertes Selbstkonzept eigener Leistungstüchtigkeit in Mathematik, attribuieren Misserfolg häufiger mit eigener Inkompetenz und glauben, Mathematik sei ein schweres Fach. Tiedemann weist aber auch darauf hin, dass dies in den ersten beiden Schuljahren noch nicht so sein müsse. Er verweist auf die Studie von Marsh et al. (1989) die für Zweitklässlerinnen eine höhere Selbsteinschätzung in Mathematik fand, als für die Knaben. Dieses Ergebnis steht nun eindeutig im Widerspruch zu den Befunden von Helmke (1997) der bereits zu Beginn der zweiten Klasse eine höhere Einschätzung in Mathematik bei den Knaben fand.

Im Sport fanden für den Kindergarten beispielsweise Althaus und Bühler (1999) im Kraftbereich (also dort, wo Mädchen glauben, weniger stark zu sein) eine tendenziell genauere Selbsteinschätzung bei den Mädchen. In Bezug auf Schnelligkeit und Geschicklichkeit überschätzen sich die Knaben gegenüber den Mädchen nur wenig mehr, die Unterschiede sind nicht signifikant. Wie wir aus den Befunden aus Kapitel 2.5. sahen, gibt es auch Leistungsbereiche im Sport, in denen sich Mädchen stärker überschätzen, als Knaben, insbesondere wenn es um Körperkoordination geht.

Hypothese 4: In der Einschätzung eigener Leistungsfähigkeit gibt es signifikante Geschlechtsunterschiede

4.1.: Knaben schätzen sich in Mathematik höher ein als Mädchen

4.2.: Knaben schätzen sich im Sport höher ein als Mädchen

4.3.: Knaben überschätzen sich in Mathematik und Sport mehr als Mädchen

4.4: Mädchen schätzen sich im Sport und in Mathematik genauer ein als Knaben

2.6.4 Unterschiedliche Sozialisierungsbedingungen von Stadt und Land

In einer Befragung von Kindern des ersten bis sechsten Schuljahres (Catani, 1997, unveröffentlicht) über das Freizeitverhalten und die Einschätzung eigener Fähigkeiten im Konfliktlösen (Sozialkompetenz) haben sich sogenannte „Stadt-Land-Effekte“ gezeigt: Kinder aus ländlichen Klassen verhalten sich eher anpassungsorientiert, die städtischen Kinder eher durchsetzungsorientiert. Die anpassungsorientierteren Kinder orientieren sich in der Schuleingangsphase aus dieser Grundhaltung heraus deutlich stärker an den Vorgaben der Lehrkräfte und müssen mit mehr Aufwand lernen, sich selbständig im Lernprozess zu organisie-

ren. Dass solche Orientierungen auch Auswirkungen auf die Selbsteinschätzung haben, kann nur vermutet werden.

Hypothese 5: Urbane Schulklassen haben eine höhere Einschätzung und eine grössere Überschätzung in Mathematik und Sport als rurale Klassen.

5.1. Urbane Schulklassen weisen im Durchschnitt eine höhere Einschätzung eigener Leistung in Mathematik und Sport auf als rurale Klassen.

5.2. Urbane Schulklassen weisen im Durchschnitt und in Mathematik und Sport eine höhere Überschätzung auf als rurale.

5.3 Urbane Schulklassen weisen im Durchschnitt eine tiefere Schätzgenauigkeit auf.

3. Methode

3.1 Die Stichprobe der Längsschnittuntersuchung

Die Auswahl der Stichprobe soll einerseits repräsentativ sein, andererseits den Anforderungen einer Längsschnittuntersuchung genügen. Ein wesentlicher Punkt ist dabei die Konstanzhaltung der Bezugsgruppe. Um zu vermeiden, dass sich vor allem im Übergang vom Kindergarten zur Schule grosse Veränderungen der Zusammensetzung ergeben, ist eine Überprüfung der Auswahl notwendig, denn Kinder gehen sehr oft in Kindergärten, die sehr nahe bei den Wohnquartieren sind: Nach einem oder zwei Kindergartenjahren werden sie dann einem grösseren Schulhaus zugewiesen und, um vergleichbare Schulklassengrößen zu erreichen, häufig neu aufgeteilt. Oder Kindergartenklassen werden aus organisatorischen Gründen aufgelöst und die Kinder verschiedenen Schulhäusern zugewiesen. In der Untersuchung von Helmke (1991) besuchten beispielsweise aus solchen Gründen durchschnittlich nur noch vier Kindergartenkinder zusammen die gleiche erste Klasse. Leicht können so Kinder beim Schuleintritt in leistungsschwächere oder –stärkere Klassen kommen, in denen höhere oder tiefere Anforderungen gelten. Die Beurteilung der eigenen Leistung kann dann anders sein. Bleiben die Klassen möglichst unverändert, dann können soziodemographische Veränderungen (beispielsweise die Einteilung der fremdsprachigen Kinder in spezielle Klassen) oder Veränderungen in der klassendurchschnittlichen kognitiven Leistungsfähigkeit (durch Weg- oder Zuzug von leistungsstärkeren oder –schwächeren Kindern) konstant gehalten werden. In jedem Kindergarten hat es aber einen Anteil von Kindern, die zurückgestellt werden oder die in eine Spezialklasse eingewiesen werden. Dann treten wiederum im zweiten Schuljahr Kinder aus den Einführungsklassen in normale zweite Klassen über.

Um Fluktuationen möglichst tief zu halten, wurden nur Klassen in die Stichprobe aufgenommen, bei denen zu erwarten war, dass 80% der Kindergartenkinder zusammen die erste Klasse besuchten. Dies erforderte eine gezielte, frühzeitige Abklärung: Ende Februar 1999 (also im Verlaufe des überprüften Kindergartenjahres) war klar, welche Kinder in die Schule übertreten werden. Aufgrund dieser Zahlen stellten die Schulleitungen dann im März 1999 die künftigen Schulklassen zusammen. Bei der Anfrage der Kindergärtnerinnen wurde dieser Sachverhalt besprochen. Aus diesem Grunde konnten dann zwei zur Mitarbeit bereite Kindergartenklassen nicht mitmachen, weil die Kinder mit vielen anderen Kindern eine erste Klasse bildeten.

Schliesslich waren es 221 Kinder, die in 14 Kindergärten im Mai /Juni 1999 untersucht wurden. Durch Wegzüge oder Repetitionen im Verlaufe der nächsten zwei Jahre gin-

gen nochmals 21 Kinder verloren, sodass für die Mathematikaufgabe die Daten von insgesamt 200 Kindern vollständig vorliegen. Wegen länger dauernden Verletzungen konnten 5 Kinder im ersten oder zweiten Schuljahr in den Sportaufgaben nicht untersucht werden, für diese Aufgaben liegen die vollständigen Daten von 195 Kindern vor. Die Ausfallquote konnte damit sehr tief gehalten werden, fast 90% der Erstuntersuchten haben über den gesamten Zeitraum von zwei Jahren mitgemacht. Nur zwei Kinder aus den teilnehmenden Klassen durften nicht an der Untersuchung teilnehmen, weil es deren Eltern ausdrücklich wünschten. Untersucht wurden grundsätzlich alle Kinder einer Klasse, auch diejenigen, die dann nicht zur Stichprobe gehörten: Die Kinder freuten sich jeweils auf die Untersuchung und so hätte deren Ausschluss diskriminierend gewirkt. Die Kindergärtnerinnen wurden im März 1999 telefonisch angefragt. Nur eine der Kindergärtnerinnen lehnte ein Mitmachen ab, alle anderen sagten zu. Die Kindergärtnerinnen hatten für die Zusage das Mitmachen ihrer Kolleginnen, die dann die Klassen weiterführten, einzuholen. Dies gelang in jedem Falle.

Für die Stichprobenziehung wurden das Berner Oberland, sowie das Emmen-, Gürbe- und Aaretal berücksichtigt. Dies geschah in der Absicht, die Studie im Einzugsgebiet des künftigen Lehrerinnenbildungsinstituts Spiez durchzuführen. In der Agglomeration Thun (Klassen Spiez und Allmendingen) leben gut 120 000 Personen, in der Agglomeration Interlaken (Matten und Interlaken) sind es 25 000 Personen. Die Gemeinde Frutigen zählt circa 5000 Personen, Brienz und Riggisberg, als kleine regionale Zentren, 2000 – 3000 Personen. Die restlichen drei Klassen stammen aus Ortschaften mit weniger als 1000 Personen. Zur Längsschnittstichprobe gehören Kindergärten, respektive Schulklassen, aus den Gemeinden, die in der folgenden Tabelle dargestellt sind (Tab.3-1-1):

Tabelle 3-1-1: Die Kinder der Stichprobe des Längsschnitts

Legende:

Total: Anzahl Kinder, für die vollständige Daten vorliegen;

Total im KG: Anzahl Kinder, die an der ersten Untersuchungswelle teilgenommen haben;

Einsch.unterhalb: Anzahl Kinder, die sich in Mathematik im Kindergarten sehr tief eingeschätzt haben, (Treppenwert –6 oder tiefer)

Einsch. oberhalb: Kinder, die sich sehr hoch eingeschätzt haben (Treppenwert +6 oder höher)

Politische Gemeinde	Soziodemografische Zuordnung	Weiblich	männlich	Total	Total Im KG	Einsch. unterhalb	Einsch. oberhalb
Riggisberg	Sehr ländlich	8	8	16	17	1	1
Mühlethurnen	Ländlich	11	8	19	21	0	1
Signau	Ländlich	6	5	11	11	2	1
Frutigen	Ländlich	16	5	21	22	0	1
Därstetten	Sehr ländlich	4	6	10	10	0	1
Brienz, Kienholz	Sehr ländlich	5	9	14	15	2	1

Brienz, Hobacher	Sehr ländlich	10	5	15	19	1	0
Interlaken West	Mischquartier	5	5	10	12	1	0
Interlaken Ost	Mischquartier	4	8	12	13	0	0
Matten	Mischquartier	6	11	17	20	1	0
Spiez, Einigen	Mischquartier	4	10	14	14	1	0
Spiezwiler	Mischquartier	8	5	13	16	1	0
Thun, Allmendingen, 1	Neuzuzüger	6	8	14	16	1	1
Thun, Allmendingen, 2	Neuzuzüger	5	9	14	15	2	0

Damit fehlen die Klassen aus ganz grossen schweizerischen Agglomerationen. Weil aber in solchen Agglomerationen grössere soziodemografische Unterschiede von Quartier zu Quartier auftreten können, müssten viele Klassen aufgenommen werden, um Klumpenrisiken zu vermeiden, die zu einem nicht-repräsentativen, städtischen Durchschnitt führen könnten. In Anlehnung an Gächter (1988, S.68f), der für die Stadt Bern eine soziodemografische Analyse gemacht hat, können die städtischen Klassen der vorliegenden Stichprobe den beiden Typen „Mischquartier“ und „Neuzuzügerquartier“ zugeteilt werden. Ein „Mischquartier“ zeichnet sich durch eine durchschnittliche schweizerische soziodemografische Verteilung aus, während in einem Neubauquartier der Ausländeranteil meist etwas tiefer ist. Die sieben Klassen der Agglomerationen Thun und Interlaken sollten also kaum soziodemografische Klumpenrisiken darstellen und für die urbane schweizerische Bevölkerung trotzdem repräsentativ sein.

Der Anteil ausländischer Kinder liegt in allen 14 Klassen zwischen circa 10% und 25%. Nur in Interlaken-West waren es 5 von 10 Kindern, welche die gesamte Untersuchung mitmachten und ausländischer Herkunft waren. Im Kanton Bern betrug die Quote ausländischer Kinder im Jahr 2000 gemäss der schulstatistischen Auswertung 14.1 % (Erziehungsdirektion des Kantons Bern, 2003), die 14 Klassen spiegeln diese Quote gut. Neben den 94 urbanen Kindern der Stichprobe liegen die Daten von 106 Kindern aus dem ländlichen Raum vor. Damit sind die ruralen Kinder mit 53 % - Anteil leicht überrepräsentiert (was, wie sich später zeigte, angesichts der meist nicht signifikanten Unterschiede zwischen diesen beiden Gruppen kaum Wirkungen auf das Gesamtergebnis hatte). 98 Mädchen stehen 102 Knaben gegenüber, hier ist die Verteilung praktisch repräsentativ: Der Mädchenanteil beträgt 49%, im Kanton Bern lag er im Jahr 2000 (Erhebung im ersten Schuljahr) bei 48.8% (Erziehungsdirektion des Kantons Bern, 2003).

Für die Stichprobenziehung der ländlichen Klassen wurden folgende weitere Faktoren berücksichtigt: Wenn die Eltern der Kinder zwar im ländlichen Raum wohnen und leben, aber in der Stadt arbeiten, dann handelt es sich häufig um Eltern, die in einer Agglomeration aufgewachsen sind und eben auch weiterhin mit dieser beruflich, aber meist auch privat, in re-

gelmässigem Kontakt stehen. Klassen mit mehr als 25% solchen Kindern sind als „ländlich“ bezeichnet (51 Vpn). Dann sind Klassen, in denen die Eltern der Kinder am Ort oder in der ländlichen Region ihr Auskommen finden, als „sehr ländlich“ bezeichnet (55 Vpn). Diese Familien haben wenige Kontakte zu urbanen Zentren (sie besuchen beispielsweise einmal monatlich ein Einkaufszentrum am Agglomerationsrand von Thun oder Bern).

Die ausgewählten Klassen sind ab erstem Schuljahr Jahrgangsklassen. Nicht berücksichtigt werden konnten also Mehrstufenklassen, in denen mehr als zwei Ausbildungsjahre zusammen eine Schulklasse besuchen. In solche Klassen werden meist pro Jahr nur wenige neue Kinder aus einem Teilzeitindergarten aufgenommen. Das führt zu so verschiedenen Bezugsgruppen in jedem Messjahr, dass auf solche Klassen verzichtet wurde. Mehrstufenklassen finden sich in städtischen und ländlichen Gebieten, weshalb sie soziodemografisch durchschnittlich nicht ländlicher als unsere rurale Stichprobe sind.

Da Stadt- und Landkinder und die Geschlechter in der Normalpopulation fast gleich verteilt sind, kann die vorliegende Stichprobe als proportional geschichtete Stichprobe bezeichnet werden. Sie erfüllt damit die Anforderungen an eine repräsentative Stichprobe mit einer Ausnahme: lernbehinderte Kinder und Kinder, die nach dem Kindergarten nicht direkt in die erste Klasse übertreten oder die das erste Schuljahr wiederholen müssen, fehlen. Dies ist ein nicht unbeträchtlicher Anteil. Etwa 5 % aller Kinder treten gar nicht in die öffentliche Volksschule ein, besuchen eine spezielle Einführungsklasse oder müssen das erste Schuljahr wiederholen. Insgesamt sind das 10% der Kinder, die in der Schuleingangsstufe einen speziellen Bildungsweg machen (Erziehungsdirektion des Kantons Bern, 2003) Eingeschlossen im Längsschnitt bleiben Kinder, die ein zweites Jahr im Kindergarten verbleiben, um einen besseren Schulstart zu haben. Solche Kinder finden sich im Wesentlichen in den ländlichen Klassen, weil in ruralen Gebieten kaum Spezialklassen in erreichbarer Nähe geführt werden.

3.2 Der Vergleich von Längsschnitt- mit Kontrollgruppen

Längsschnittuntersuchungen bestehen aus Messungen von Variablen über eine Folge von Messzeitpunkten. Auch wenn diese in einem grösseren zeitlichen Abstand erfolgen, so sind Effekte, die durch die Messwiederholung entstehen, nicht ganz auszuschliessen. Beim zweiten Messzeitpunkt könnten Erinnerungs- und Trainingseffekte die Befunde beeinflussen und damit die interne Validität beeinträchtigen. Deshalb wurden Mittelwerte, Varianzen und Korrelationen der Längsschnittgruppe mit denjenigen je einer Kontrollgruppe (KG1 und KG2) verglichen: Im ersten und zweiten Schuljahr wurden zusätzlich Daten bei je rund

hundert Kindern erhoben, die nur einmal an der Untersuchung teilnahmen, entweder als Erst- oder Zweitklässler. Die KG 1 (Erstklässler) besteht aus 45 urbanen 56 ruralen Kindern, in der KG 2 (Zweitklässler) sind es 50 urbane und 48 rurale Kinder. Beim Geschlecht sind die Knaben in beiden Kontrollgruppen leicht, aber nicht signifikant, übervertreten. Alle Kontrollgruppenkinder stammen aus der gleichen Stichprobengesamtheit wie die Experimentalgruppe, aus dem Emmental und aus dem Berner Oberland.

Mögliche Kohorteneffekte, die die interne Validität beeinträchtigen könnten, sind demgegenüber vermutlich gering, denn über den Messzeitraum hat sich an den schulischen Rahmenbedingungen wenig verändert: die Lehrmittel sind die gleichen geblieben, der gültige Lehrplan war bei der Erstuntersuchung seit 4 Jahren eingeführt. Grössere Reformen wurden im Untersuchungszeitraum an den untersuchten Schulen ebenfalls nicht durchgeführt. Das Lebensumfeld der Kinder war vergleichbar, die Arbeitslosigkeit der Erwachsenen, die sicher auch die Familien einiger Kinder betraf, bewegte sich in diesem Zeitraum zwischen gesamtschweizerisch 50 000 und 100 000 betroffenen Personen. Im Untersuchungsgebiet gab es über diesen Zeitraum keine grösseren Firmenschliessungen oder Arbeitsplatzreduktionen, die Familien von untersuchten Kindern betroffen hätten. Solche Veränderungen waren bereits zu Beginn der neunziger Jahre mit der Schliessung einer grossen Giesserei in Uetendorf bei Thun abgeschlossen. Nun die Befunde zu den Signifikanztests Kontroll- versus Längsschnittstichprobe (Tab.3-2-1).

Tabelle 3-2-1: Vergleich von Längsschnitt- (LG) und Kontrollgruppe 1 (KG1) im ersten Schuljahr
SD: Standardabweichung; M: Mittelwert

	SD LG	SD KG1	F	Sig. F	DF	t-Wert	Sig. t	M LG	M KG1
Mathematik Einschätzung	2.59	2.75	.170	.681	309	-.1600	.973	1.54	1.59
Mathematik Leistung	2.39	2.24	.153	.696	309	2.118	.035*	24.65	24.05
Hüpfen (1) Einsch. Treppe	2.40	2.73	4.26	.040*	174	-1.498	.136	.68	1.16
Hüpfen (1) E. an Aufg.	2.42	1.66	5.99	.015*	167	-2.169	.032*	6.12	6.54
Hüpfen Leistung	1.78	1.68	.400	.528	308	-.120	.905	6.64	6.67
Balancieren Einsch. Treppe	2.78	2.83	.189	.664	308	-1.316	.189	.27	.71
Balancieren E. an Aufg.	81.22	81.17	.093	.761	308	-1.315	.190	224	237
Balancieren Leistung	81.14	74.12	1.74	.187	308	-.409	.683	134	138

((1) : Berechnung nach Ward für ungleiche Varianzen; Mittelwerte nach verschiedenen Masseinheiten, vgl. Text)

Die F-Werte zeigen an, ob zwischen Kontroll- und Längsschnittgruppe signifikante Varianzunterschiede bestehen, in der ersten Klasse trifft dies auf die Einschätzung der Hüpf-aufgabe zu, die Einschätzungen der Kontrollgruppe 1 streuen signifikant mehr. Zur Überprüfung der Mittelwerte wird deshalb das Verfahren von Ward für einen Vergleich von Stichproben mit ungleichen Varianzen zurückgegriffen. Bei der Leistung sind die Streuungen in der Längsschnittgruppe höher. Die Einschätzungen der eigenen Leistungen weichen in den Mittelwerten zwischen den beiden Gruppen in keinem Fall signifikant ab, die Unterschiede bewegen sich innerhalb der Zufallschwankungen. Die Kinder der Kontrollgruppen haben sich auf der Treppenstufenskala in Mathematik und Hüpfen nur geringfügig höher eingeschätzt, im Balancieren beträgt der Unterschied 0.46 Treppenstufen. In der Leistungsfähigkeit unterscheiden sich die Kinder in der Mathematik signifikant: die Erstklässler der Längsschnittgruppe waren im KAB-C 0.60 Rohwertpunkte besser. Im Gegensatz dazu waren die Kinder der Kontrollgruppe 1 in den Sportaufgaben minim besser, sie haben sich an der Aufgabe entsprechend etwas besser eingeschätzt.

Die Schätzgenauigkeit (Einschätzung an der Treppe x Leistung) beträgt in der Kontrollgruppe 1 $r=0.041$, in der Längsschnittstudie ist sie höher ($r=0.181$), die Differenz ist nicht signifikant ($z=1.109$; $p> 0.10$).

In Bezug auf die abhängige Variable erfüllt die interne Validität der zweiten Erhebung die Erwartungen, die Unterschiede in den Befunden schwanken zwischen der Kontrollgruppe 1 und der Längsschnittgruppe nur im Zufallsbereich, bei der Leistung (unabhängige Variable) ergibt sich eine signifikante Differenz, die Kinder der Kontrollgruppe 1 sind leistungsschwächer.

Nun zum zweiten Schuljahr (Tab.3-2-2):

Tabelle 3-2-2: Vergleich von Längsschnitt- und Kontrollgruppe 2 im zweiten Schuljahr

	SD LG	SD KG2	F	Sig. F	DF	t-Wert	Sig. t	M Exp	M Kontr.
Mathematik Einschätzung	2.40	2.24	.051	.821	291	1.718	.087	1.15	0.64
Mathematik Leistung	2.85	2.53	.035	.853	291	3.442	.001**	27.20	25.82
Hüpfen (1) Einsch. Treppe	1.98	1.25	.348	.555	291	2.155	.032*	.79	.23
Hüpfen (1) E. an Aufg.	1.56	1.28	3.143	.077	290	-.521	.603	7.09	7.19
Hüpfen Leistung	1.75	1.79	.080	.778	290	.466	.642	7.84	7.74
Balancieren Einsch. Treppe	2.18	2.62	2.258	.134	291	1.883	.061	.18	-.36
Balancieren E. an Aufg.	66.37	73.95	.750	.387	291	-1.389	.166	203	214
Balancieren Leistung	99.63	88.22	2.417	.121	291	2.905	.004**	194	159

(1) : Berechnung nach Ward für ungleiche Varianzen; Mittelwerte nach verschiedenen Masseinheiten, vgl. Text)

Die Einschätzungen der Leistungen anhand der Treppenstufenskala weichen in Mathematik und Balancieren nicht signifikant ab. Im Gegensatz zur Kontrolluntersuchung in der ersten Klasse liegen die Einschätzungen durch die Kontrollgruppe 2 hier aber unterhalb der Längsschnittgruppe. Dies gilt auch für das Hüpfen: In dieser Aufgabe schätzen sich die Kinder aber signifikant tiefer ein.

Die Leistungen liegen bei der Kontrollgruppe 2 tiefer. Die Kinder der Längsschnittgruppe haben bei der dritten Durchführung den Test bereits zum dritten Mal durchgeführt, was ihre Leistungsfähigkeit allenfalls erhöht haben könnte. Dies war insbesondere in Mathematik nicht erwartet worden, denn die Re-Test-Reliabilität ist im KAB-C sehr hoch: Innerhalb 2 Wochen stieg die Leistung im Re-Test nur sehr wenig an (nicht signifikant). Im Vergleich zu den KAB-C Normen sind die Leistungen der Längsschnittgruppe exakt so hoch wie durchschnittlich erwartet, nämlich bei 27.2 Rohwertpunkten. Die Kontrollgruppe 2 demgegenüber erbrachte signifikant tiefere Leistungen ($t(307) = 2.118$, $p < 0.05$). Wie ist dieser Leistungsunterschied erklärbar? In die zweiten Klassen treten jeweils am Ende des ersten Schuljahres Kinder ein, die aus Einführungsklassen stammen, in denen der Lernstoff des ersten Schuljahres in zwei Jahren bewältigt wird, dies trifft insbesondere auf die Klassen aus den städtischen Gebieten zu. In den ländlichen Klassen gibt es dagegen eher Wiederholungen des Kindergartenjahres. Im Längsschnitt fehlen in den Befunden der Kindergartenklassen diejenigen, die den Kindergarten wiederholen, weil sie ja dann ein Jahr später nicht an

der zweiten Messung teilnehmen. Eine Prüfung ergab nun, dass in der Kontrollgruppe 2 zehn von 63 städtischen Kindern aus Spezialklassen in die entsprechenden Kontrollklassen übergetreten sind. Einerseits sind diese Kinder ein Jahr älter als die anderen, andererseits waren sie im Kindergarten in ihrer sozialen Entwicklung noch im Rückstand oder sie hatten Sprachdefizite, die nun in der zweiten Klasse ausgeglichen sein sollten. Auf jeden Fall befinden sich in der Längsschnittgruppe keine vergleichbaren Kinder. Aus diesem Grunde wurde eine weitere Kontrolluntersuchung vorgenommen, ein Vergleich ohne die Kinder, die repetiert haben (Tab.3-2-3).

Tabelle 3-2-3: Vergleich von Längsschnitt- und Kontrollgruppe 3 (n= 86) im zweiten Schuljahr, ohne Repetenten

	SD LG	SD KG2	F	Sig. F	DF	t-Wert	Sig. t	M Exp	M Kontr.
Mathematik Einschätzung	2.40	1.96	1.26	.290	280	1.616	.110	1.15	0.67
Mathematik Leistung	2.85	2.34	4.79	.029*	189	2.58	.011*	27.20	26.36
Hüpfen (1) Einsch. Treppe	1.98	2.02	1.04	.309	280	2.94	.004**	.79	.02
Hüpfen (1) E. an Aufg.	1.56	1.25	4.54	.034*	189	-.019	.985	7.09	7.10
Hüpfen Leistung	1.75	1.82	.157	.692	280	.862	.390	7.84	7.64
Balancieren Einsch. Treppe	2.18	2.42	.285	.594	280	1.632	.10	.18	-.30
Balancieren E. an Aufg.	66.37	70.91	.008	.928	291	-.744	.458	203	210
Balancieren Leistung	99.63	88.50	1.74	.188	291	3.101	.003**	194	155

((1) : Berechnung nach Ward für ungleiche Varianzen;

Mittelwerte nach verschiedenen Masseinheiten, vgl. Text)

Die Unterschiede sind tatsächlich kleiner geworden, aber die Unterschiede sind auch unter dieser Bedingung in den gleichen Variablen signifikant. Neben den Mittelwerten und Varianzen wurde noch die Genauigkeit von Einschätzung x Leistung kontrolliert. Sie ergibt in Mathematik für die Kontrollgruppe 2 den Wert $r = 0.386$, für die Längsschnittgruppe $r = 0.348$, der Unterschied ist nicht signifikant ($z = 0.349$; $p > 0.10$).

Zusammenfassend kann gesagt werden, dass die externe Validität aufgrund des gewählten Vorgehens hoch ist, die Befunde können auf andere Kinder generalisiert werden. Die interne Validität ist in Bezug auf die Einschätzungen der Leistungen gegeben. In Bezug

auf die unabhängige Variable „Leistung“ im zweiten Schuljahr, also bei der dritten Erhebung, ist die interne Validität teilweise eingeschränkt: In Mathematik erbrachte die Kontrollgruppe eine tiefere als vom KAB-C erwartete Durchschnittsleistung, während die Längsschnittgruppe exakt im Leistungsdurchschnitt ist. Im Sport sind die Leistungen der Kontrollgruppe im Balancieren ebenfalls tiefer, nicht jedoch im Hüpfen

3.3 Vorgehen

Im Dezember 1998 wurde die Erziehungsdirektion des Kantons Bern schriftlich angefragt, ob die Untersuchung die notwendigen Voraussetzungen erfülle, um an öffentlichen Schulen durchgeführt zu werden. Nach dem Vorliegen der Bewilligung wurden Schulleitungen um Adressen der Kindergärtnerinnen angefragt. Anschliessend erfolgte die telefonische Anfrage bei den Kindergärtnerinnen, Ziel und Art der Erhebung wurden geschildert. Das nötige Mitmachen der Unterstufenlehrkräfte ein und zwei Jahre später wurde ebenfalls thematisiert. Die Kindergärtnerinnen des Berner Oberlandes, die nur eher selten an Untersuchungen mitmachen können, zeigten sich sehr interessiert. Nur eine Kindergärtnerin konnte nicht mitmachen, da sie im fraglichen Zeitraum bereits viele Aktivitäten geplant hatte. Nachdem die Kindergärtnerinnen ihre Kolleginnen an den zukünftigen ersten Klassen informiert hatten, sagten alle 16 angefragten Kindergärtnerinnen zu. Nachdem geklärt war, wieviele Kinder des betreffenden Kindergartens zusammen die erste Klasse besuchen werden, konnten 14 Klassen im April 1999 definitiv in die Untersuchung aufgenommen werden. Anschliessend wurde das Einverständnis der Schulkommissionen eingeholt. Dann wurden die Eltern mittels Elternbrief über das Vorhaben und dessen Ziele informiert. Sie konnten schriftlich von ihrem Recht Gebrauch machen, ihr Kind von der Untersuchung zu dispensieren. Dies geschah in zwei Fällen. Im April wurde bei zwei Kindergärten eine Pilotierung vorgenommen. Das Ergebnis entsprach den Erwartungen, Veränderungen mussten keine mehr vorgenommen werden. Die Erhebung selbst wurde von Lehrerstudentinnen und -studenten des Seminars Spiez durchgeführt, die kurz vor dem Ausbildungsabschluss standen. Mittels eines Instruktionsvideos konnten sie den Gesamttablauf kennen lernen. In zwei Sitzungen wurde die Instruktion der Erhebung sowie die konkrete Durchführung mit den Instrumenten geübt (vgl. das Erhebungsmanual im Anhang). Pro Klasse waren vier Studierende gleichzeitig im Einsatz, die am ersten Erhebungstag zusätzlich durch Assistenten des Psychologischen Institutes vor Ort gecoacht wurden. Im Juni 1999 wurde die erste Erhebungswelle ohne Zwischenfälle durchgeführt, die Durchführungszeitspanne (Tage vom ersten zum letzten Erhebungstag) betrug zwei Wochen

Im November 1999 wurden die Kindergärtnerinnen sowie die Lehrkräfte, die an der sechs Monate später stattfindenden zweiten Welle mitmachten zu einer ersten Information eingeladen. Im Zentrum standen Informationen zur Theorielage und den bisherigen Befunden. Präsentiert wurden Mittelwerte für die Gesamtgruppe aus der ersten Erhebungswelle, die aber keine Rückschlüsse auf die eigene Klasse zuließen. Im Winter wurden zusätzliche Pilotierungen für die Aufgabe „rückwärts Balancieren“ vorgenommen. In der Durchführung schnitten die Kinder der zweiten Untersuchungswelle dann überraschenderweise im Durchschnitt schlechter ab als von den Pilotierungen her erwartet wurde, während sie in der dritten Welle die erwartete Durchschnittsleistung erbrachten.

Die zweite und dritte Erhebungswelle erfolgte fast im Jahresabstand, ab Mitte Mai bis im Juni 2000 und 2001. Die Durchführungszeitspanne war wegen der Feiertage etwas grösser, nämlich drei Wochen. Für die zweite und dritte Welle mussten neue Studierende eingesetzt werden. Sie wurden nach dem gleichen Verfahren geschult und gecoach, wie im ersten Erhebungsjahr. Der Messzeitpunkt lag in allen Jahren 3 bis 4 Wochen vor dem Ende des Ausbildungsjahres. Im Kindergarten hatten die Kinder anhand des Schuleintrittsgesprächs bereits im Februar / März eine Leistungsrückmeldung erhalten, die, gemäss den kantonalen Vorgaben, den allgemeinen kognitiven und motorischen Entwicklungsstand festhielt. Sie wussten zudem seit März, ob sie nun den Kindergarten ein weiteres Jahr besuchen würden oder ob sie im kommenden August in die Schule gehen. Die Erstklässler hatten im Februar / März ebenfalls eine erste Leistungsrückmeldung erhalten, die die Erreichung der Lernziele und den gemachten Lernfortschritt in Mathematik und Sport schriftlich festhielten. In der zweiten Klasse war das Verfahren gleich, die Kinder hatten aber zusätzlich Ende des ersten Schuljahres ein erstes Zeugnis erhalten. Dieses beurteilte die Leistungen mit den Begriffen „nicht erreicht“, „teilweise erreicht“, „erreicht“ oder „übertroffen“ summativ pro Fach und gemessen an den angestrebten Lernzielen. Zwischen diesen direkten Leistungsrückmeldungen und dem Messzeitpunkt lagen in allen Fällen drei Wochen Ferien und sechs Wochen zusätzliche Unterrichtserfahrung. Es wurde auch darauf geachtet, dass die Abgabe des Zeugnisses in den beiden Schuljahren nicht unmittelbar bevorstand.

3.4 Erhebungsinstrumente

3.4.1 Leistungstests in Mathematik

Die Kindergartenkinder machen in ihrem privaten Alltag Erfahrungen im Zahlenraum und können meist einfache Additionen und Subtraktionen im Zahlenraum von 1 bis 10 vornehmen. In der deutschschweizerischen Kindertadition gehört das gezielte Schulen der Rechenfertigkeiten nicht zum Ausbildungsplan (wie es etwa in den französischsprachigen Landesteilen der Fall ist). So findet sich im ersten Lehrplan für den Kindergarten des Kantons Bern (Erziehungsdirektion des Kantons Bern, 1999) nur ein gezielter Hinweis zur Förderung der mathematischen Fähigkeiten: „Mathematische Gesetzmässigkeiten kennenlernen wie teilen, Gruppen bilden, addieren und subtrahieren“ (Erziehungsdirektion des Kantons Bern, 1999, 36). Der deutschschweizerische Kindergarten ähnelt damit demjenigen in Deutschland, hingegen ist es in den meisten europäischen Ländern und den USA üblich, dass Kinder bereits im Kindergartenalter systematisch in ihren mathematischen Fähigkeiten beschult werden. Damit ergibt sich das Problem, dass eine Erhebung der mathematischen Fähigkeiten bei Kindergartenkindern davon auszugehen hat, dass Kinder in diesem Alter allenfalls private Erfahrungen mit dem gezielten Lösen von solchen Aufgaben haben. Entsprechend müssen die Aufgaben altersadäquat, sehr lebensnah und anschaulich sein.

Bei der Auswahl eines Instrumentes zur Messung der Leistung in Mathematik wurden die fachlichen Kenntnisse der Erziehungsberatungsstellen einbezogen, die über langjährige testdiagnostische Erfahrungen im Schuleingangsalter verfügen. Eingesetzt werden von diesen Stellen im Wesentlichen der Intelligenztest von Kramer (1972), der neben Anderem auch Rechenaufgaben enthält. Dann kommt ein Rechentest, der auf die deutschschweizerischen Lehrmittel der späten siebziger und achtziger Jahre adaptiert ist, zum Einsatz (Lobeck, 1987), sowie, seit wenigen Jahren, die „Kaufman-Assessment-Battery for Children“ in einer deutschsprachigen Fassung (Melchers & Preuss, 1994).

Der Test von Kramer (1972) ist auf Lehrmittel abgestützt, die seit Mitte der siebziger Jahre nicht mehr eingesetzt werden, die Normierung beruht auf einer Schweizer Stichprobe. Er kann sowohl im Kindergarten, wie in den ersten Schuljahren eingesetzt werden, würde sich also für die Anwendung in allen Messjahren eignen. Dem heutigen Lernstoff des ersten und zweiten Schuljahres entspricht er aber kaum mehr, was die Validität deutlich einschränkt.

Aufgebaut auf den bis Ende der neunziger Jahren weit verbreiteten Lehrmitteln ist der Leistungstest von Lobeck (1987) welcher ebenfalls auf einer schweizerischen Normstichprobe beruht. Allerdings gibt es keine Ausgabe für den Kindergarten, sodass er nicht durchgän-

gig eingesetzt werden könnte. Die weitgehende Konstanthaltung des Messinstrumentes war aber eine zentrale Forderung an die Untersuchung, weil bisherige Untersuchungen die Leistung der Kinder im Kindergartenalter und der ersten Klasse häufig über die Befragung der Lehrkräfte oder der Eltern erhoben (z. B. Helmke, 1991). In der zweiten Klasse kam es dann zu einem Wechsel des Messinstrumentes, was den Vergleich der Einschätzgenauigkeit zwischen den Messzeitpunkten deutlich erschwerte. Amerikanische Untersuchungen verwenden zwar häufig konstante Leistungstests, aber dort gibt es für uns zwei Einschränkungen. Erstens finden bereits im Kindergartenalter Mathematiksequenzen im Unterricht statt und zweitens sind diese Kinder durchschnittlich ein Jahr jünger als die schweizerischen.

Weil es von Interesse ist zu untersuchen, wie sich Kinder in der Höhe der Selbsteinschätzung und in der Schätzgenauigkeit unterscheiden, je nach dem ob sie Mathematikunterricht haben oder nicht, kam deshalb nur ein Testverfahren in Frage, das diesen Anforderungen möglichst nahe kommt. Die auf dem amerikanischen KAB-C aufbauende deutschsprachige Fassung von Melchers und Preuss (1994) besteht aus 16 Untertests, von denen der Fertigkeitstest „Rechnen“ ab dem vierten Lebensjahr eingesetzt werden kann. Mit diesem Untertest können die Leistungen in Mathematik erhoben werden. Er beruht auf einer Normierungsstichprobe von 3098 Kindern aus dem deutschsprachigen Raum (erhoben zwischen 1986 und 1989), von denen 6.9% aus der Schweiz stammen. Für den Kindergarten, die erste und zweite Klasse sind das zwischen 322 und 349 Kinder, allerdings stammen im Mittel nur 24 Kinder pro Jahrgang aus der Schweiz. Bei den Kindern der Normstichprobe, die das Alter unserer Kindergartenkinder haben, sind 227 tatsächlich im Kindergarten, aber 87 gehen bereits zur Schule. Die Kinder der Normstichprobe könnten deshalb leistungsmässig im ersten Messjahr etwas höher liegen, als die Kinder der vorliegenden Untersuchung, die ja alle im Kindergarten sind. Es besteht also ein gewisses Risiko, dass die Kinder der vorliegenden Untersuchung im Durchschnitt von der Normstichprobe abweichende Ergebnisse erzielen. Dieses Risiko wurde aber aus drei Gründen in Kauf genommen: Erstens sind die Mathematikaufgaben in einer Rahmengeschichte mit Kindern eingebaut und alle Aufgaben werden anhand eines farbigen Bildes dargeboten, was auch für Kinder, die noch nicht in einem Lehrmittel zu arbeiten gewohnt sind (Kindergarten) eine angemessene Form der Aufgabendarbietung darstellt. Zweitens ist die curriculare Validität der Aufgaben gegenüber den anderen Testverfahren höher, weil die gemessenen Fähigkeiten im Bereich der Mathematik (das Wissen über Zahlen und einfache mathematische Konzepte, vgl. das Handbuch zum KAB-C, S.251ff) besser mit denjenigen der in der Schule zur Anwendung gelangenden Lehrmittel übereinstimmt, als bei anderen Testverfahren. Die Lern- und Leistungsziele des Lehrmittels „Zahlenbuch“, welches in 12 der 14 Klassen im ersten und zweiten Schuljahr zum Einsatz gelangt, stimmen sehr gut mit den überprüften Fähigkeiten im KAB-C überein, während der Leistungstest von Lobeck (1987) optimal auf die Lernziele des Lehrmittels der

beiden restlichen Klassen abgestimmt ist. Dieses ältere Lehrmittel enthält aber ähnliche Lernziele und diente als Grundlage zur Entwicklung des „Zahlenbuch“, sodass diese beiden Klassen in der Stichprobe verblieben. Zum Dritten ist die Durchführungsobjektivität beim KAB-C deutlich höher als bei anderen Testverfahren. Die bei der Erhebung eingesetzten Studierenden kennen das Lehrmittel „Zahlenbuch“ aus eigener Erfahrung im Unterrichten. Die Mathematikaufgaben des KAB-C sind vergleichbar konstruiert, während die Aufgaben aus den anderen Testverfahren den Einsatz von Spezialisten voraussetzen. Ein Vorteil der KAB-C Aufgaben ist auch darin zu sehen, dass fremdsprachige Kinder die Aufgaben gut verstehen, weil Bilder die Berechnungsgrundlagen abbilden: In der Pilotierung wurde der Überprüfung dieses Aspekts stark Rechnung getragen, weil der Anteil fremdsprachiger Kinder ja nicht unbedeutend ist und weil diese im Schuleingangsalter deutlich stärkere sprachliche Defizite aufweisen, als in höheren Schuljahren. Insgesamt wurde nur in zwei Fällen durch Studierende von Kindern berichtet, die vermutlich die Aufgaben nicht vollständig verstanden haben.

Neben der Objektivität und der Validität des Messinstrumentes zur Erhebung der Mathematikleistungen im Schuleingangsalter interessiert die Reliabilität. Die „split-half“-Reliabilität des KAB-C ist in Mathematik im Schuleingangsalter sehr hoch (0.84 bis 0.92), bei einer Testwiederholung innert 19 Tagen stieg der Mittelwert bloss von 93,3 (nicht signifikant) auf 93,4 Rohwertpunkte. Die Korrelation der Messungen beträgt $r=0.90$. Die Messgenauigkeit kann also als hoch bezeichnet werden.

Beim KAB-C handelt es sich um ein adaptives Verfahren, das heisst, es werden nicht alle Aufgaben angeboten, sondern ein altersabhängige Auswahl. Der Untertest „Rechnen“ besteht aus 37 Aufgaben. Im Kindergarten beginnt man mit Aufgabe 6, in der ersten Klasse mit der Aufgabe 15 und ein Jahr später mit der Aufgabe 19. Immer vier oder fünf Aufgaben gehören zusammen. Wenn von diesen Aufgaben nicht zumindest eine richtig gelöst ist, kann nicht zur nächsten Gruppe weitergegangen werden. Falls aus der zu Beginn gestellten Aufgabengruppe kein richtiges Ergebnis resultiert, muss weiter zurückgegangen werden. Dies war in der vorliegenden Untersuchung nie der Fall, die tiefste Leistung lag bei 8 Rohwertpunkten, die höchste bei 36, also nur 3 Punkte unter dem Testmaximum: Ein Zweitklässler erreichte damit eine Leistung, die derjenigen eines guten Schülers des sechsten Schuljahres entspricht. Vorteile dieses Testverfahrens sind die in jedem Jahr ungefähr gleiche Testdauer, die an die Motivation und Konzentration vergleichbare Anforderungen stellt und das Abdecken des ganzen möglichen Leistungsvermögens der Kinder, auch dann wenn dieses deutlich unter- oder oberhalb des Testalters liegt. Damit erlaubt dieser Test eine hohe Aussagefähigkeit über eine sehr grosse Leistungsbreite. Trotzdem liegen die Befunde in den beiden Schuljahren vollumfänglich innerhalb der jeweiligen Konfidenzintervalle (95%-Niveau), während im Kindergarten 3 Kinder erst beim 99% -Niveau innerhalb des Konfidenzintervalles

liegen, sie sind haben sehr tiefe Leistungen erbracht. Damit die Kinder nicht mit einem Misserfolgserlebnis aufhören mussten (das ist das Abbruchkriterium des Tests), wurden nach der letzten Aufgabe aus dem KAB-C anhand eines Bildes aus dem Test nochmals drei einfachere Aufgaben gestellt, die das Kind sicher lösen konnte. Ansonsten folgte die Untersuchung den Vorgaben des Testmanuals. Die Rechenaufgaben wurden jedem Kind einzeln gestellt, es hatte jeweils unbeschränkte Lösungszeit, es ist also ein Powertest- Verfahren. Ein Speedtest, wie er auch von Helmke ab zweiter Klasse angewandt wurde, würde dem Lernverständnis im Kindergarten nicht gerecht. Bleibt zu sagen, dass die Kinder auch bei der letzten Durchführung immer noch sehr motiviert waren, obwohl sie die Rahmengeschichte ja bereits kannten.

Die Normtabellen des KAB-C haben den Vorteil, dass sie für je drei Lebensmonate abgestuft sind und nicht ganze Lebensjahre zusammenfassen. Dies ist deshalb von Bedeutung, weil die Leistungen im Schuleingangsalter relativ schnell zunehmen und der Messzeitpunkt dann das Ergebnis erheblich beeinflussen könnte. In der Untersuchung waren die Kinder im Kindergarten durchschnittlich 6.8 Jahre alt, die entsprechenden Normen gelten für Kinder von 6.6 bis 6.8 Jahren. Ein und zwei Jahre später waren die Kinder durchschnittlich 7.7 und 8.7 Jahre alt (die Untersuchung musste gegenüber dem Kindergarten durchschnittlich drei Wochen früher durchgeführt werden).

3.4.2 Die Instrumente für den Leistungsbereich Sport

Im Leistungsbereich „Sport“ gibt es verschiedenste altersgerechte Aufgaben um jüngere Kinder mit einer ansprechenden Herausforderung zu konfrontieren. Die Operationalisierung der sportlichen Fähigkeiten erfolgt durch die Messung von Fertigkeiten in bestimmten Leistungsbereichen. Diese Fertigkeiten gehen auf zwei Gruppen von motorischen Fähigkeiten zurück, auf die konditionellen und auf die koordinativen Fähigkeiten. Zur ersten Gruppe gehören die Ausdauer-, die Kraft- und Schnelligkeitsfähigkeit, zur zweiten Gruppe Koordination, Präzision und Koordination unter Zeitdruck. (Althaus & Bühler, 1999). Für unseren Zusammenhang ist nun wichtig, dass Kinder Fertigkeitsunterschiede unter sich und Fortschritte bei sich selber wahrnehmen können, wenn sie in der Lage sein sollen ihre Leistungsfähigkeit einzuschätzen. Aus der ersten Gruppe trifft dies vor allem auf Kraftfertigkeiten zu, aus der zweiten Gruppe ist dies die Gleichgewichts-, respektive Balancierfertigkeit. Der Körperkoordinationstest (KTK) von Kiphard und Schilling (1974) umfasst vier Aufgaben, nämlich Balancieren rückwärts, monopedaless Überhüpfen, seitliches hin und her Springen, sowie seitliches Umsetzen. Aus diesen vier Aufgaben wurden das Balancieren rückwärts als Fertigkeit für die Messung der koordinativen Fertigkeiten und das Monopedaless Überhüpfen als Messung für

die Krafftähigkeit ausgewählt. Die Leistungsmittelwerte von Knaben und Mädchen sind im Balancieren rückwärts nicht signifikant unterschiedlich, gemäss den KTK-Normen sind die Mädchen in allen Jahren (aber immer weniger) besser als die Knaben. Im monopedalen Überhüpfen sind die Mädchen im Kindergartenalter signifikant besser ($p=0.001$), in der ersten Klasse dann tendenziell immer noch leicht überlegen, während in der zweiten Klasse die Knaben bessere Leistungen erbringen. Sollten sich die Knaben nun leistungsmässig höher einschätzen, so liegt das nicht an den gemachten Erfahrungen, denn Mädchen sind in diesem Alter durchwegs, wenn auch meist nicht signifikant, leistungsstärker. Die Befunde des KTK wurden ebenfalls auf Unterschiede zwischen Stadt- und Landkindern geprüft. In Balancieren rückwärts waren die Landkinder, im Monopedalen Überhüpfen die Stadtkinder leistungsstärker, die Unterschiede sind aber bei weitem nicht signifikant. Zur Reliabilitätsprüfung wurden von Kiphard und Schilling Re-Test-Analysen gemacht. Bei Testwiederholung nach vier Wochen fand sich beim Balancieren rückwärts ein $r=0.80$, beim monopedalen Überhüpfen ein $r=0.96$.

Die Aufgabe „monopedales Überhüpfen“ wurde unverändert aus dem KTK übernommen. Bei dieser Aufgabe hatten die Kinder einbeinig über ein Hindernis zu hüpfen. Dieses bestand aus einem Schaumgummi, der 50 cm breit, 20 cm tief und 5 cm hoch war. Wurde das Hindernis ohne Berührung, Sturz oder beidbeinige Landung übersprungen, galt die Höhe als erfolgreich übersprungen. Dann wurde auf den Schaumgummi ein weiteres Stück gelegt und es ging auf der neuen Höhe weiter. Pro Höhe waren maximal drei Versuche möglich. Wurde die Höhe nicht mehr erfolgreich übersprungen, dann wurde der Versuch abgebrochen. In jedem Jahr sollten mit möglichst gleich vielen Sprüngen die Leistungen erbracht werden, damit nicht die zunehmende Anzahl der Sprünge und damit die muskuläre Ermüdung, zum Misserfolg führten. Im Kindergartenjahr begann die Aufgabe auf der Höhe von 5 cm, also einem Schaumgummi, während in der zweiten Klasse mit der Höhe 20 cm gestartet wurde. Der Versuchsleiter machte die Aufgabe einmal vor. Dann hatten die Kinder einen „Testsprung“ auf der untersten Höhe, damit auch die Richtigkeit der Durchführung geprüft werden konnte. Das monopedale Überhüpfen wurde wie die Mathematik in einer adaptiven Form durchgeführt: die zunehmende Aufgabenschwierigkeit wurde durch die lineare Zunahme (immer ein Schaumgummi mehr) gewährleistet, damit die Aufgabe in jedem Jahr aber vergleichbar schwierig war, wurde auf jeweils auf einer Höhe begonnen, die 99% der Kinder bewältigen können und die zur möglichst gleichen Anzahl Sprünge führte (vgl. Kiphard & Schilling, 1974, Manual, 32). In Ergänzung zur KTK-Versuchsanordnung wurde im Landebereich eine Turnmatte hingelegt, um Fussverletzungen zu vermeiden.

Die Aufgabe „Balancieren rückwärts“ konnte nicht in der vorgegebenen Form durchgeführt werden, weil sie sehr viel Zeit pro Kind gebraucht hätte und weil für die Durchführung ein ganzer Raum nötig gewesen wäre: Es sollten nämlich drei Balken auf den Boden gelegt

werden, die je 3 Meter lang waren und unterschiedliche Breiten (6 cm, 4.5cm, 3 cm) hatten. Ein wichtiger Grund für die Abänderung ist, dass die Aufgabe in der Testform des KTK nicht adaptiv ist: In allen Jahren beginnt man am gleichen Punkt. Dadurch wird, als Teil des Balanciervermögens, die Konzentration, respektive die zunehmende Müdigkeit für den Erfolg bedeutsamer. In dieser Form ist die Aufgabe aber auch nicht sehr realitätsnah: In der Schule werden Lernaufgaben auf das Schulalter hin adaptiert, ein Training im Sport beginnt in jedem Jahr auf einer etwas anspruchsvolleren Stufe. Aus diesen Gründen wurde die Aufgabe in eine adaptive Form umgewandelt und zusätzlich vereinfacht, als Grundlage dienten die Angaben aus dem KTK, sowie Vorversuche.

Die Aufgabe bestand aus einem sechs Meter langen Balken, der mit einer Balkenbreite von 5.5 cm begann, konisch war und mit der Breite von 2.5 cm endete. Der Balken war 2 cm dick und bestand aus 1 Meter langen Stücken, die am Boden mit Klebeband fixiert wurden. Die Aufgabe wurde in einer adaptiven Form präsentiert. Im Kindergarten hatten die Kinder die ersten vier Meter vor sich, im ersten Schuljahr die Strecke von Meter 2 bis 5 (also wiederum 4 Meter Gesamtlänge) und im zweiten Schuljahr dann die Meter 3 bis 6. Aufgrund der Vorversuche und der Pilotierung an zwei Klassen im Kindergarten und der zweiten Klasse zeigte sich, dass die Kindergartenkinder durchschnittlich fast die mittlere Länge (2 Meter Distanz) erreichten, bis sie den Boden berührten, die Abweichungen waren nicht signifikant.

Die erfolgreich erreichte Länge wurde hinter dem Fuss gemessen, der als letzter den Boden noch nicht berührte. Aufgrund der Angaben aus dem KTK konnte angenommen werden, dass die Leistungszunahme vom Kindergarten zur ersten Klasse deutlich grösser ist als von der ersten zur zweiten Klasse. In der Untersuchung traf dann diese Prognose nicht ein, denn überträgt man die tatsächlich erbrachte Leistung auf die gesamten 6 Meter, dann kamen die Kinder bei der ersten Erhebung bis 164 cm, in der zweiten bis 233 cm und in der dritten bis 391 cm. Der tatsächliche Leistungszuwachs war also, wider Erwarten, vom ersten zum zweiten Schuljahr grösser. Hätte man allerdings nach den Pilotierungen die Aufgabe abgeändert, wäre die gleichmässige Zunahme der Aufgabenschwierigkeit nicht möglich gewesen. Es wurde auf eine Abänderung verzichtet, nicht zuletzt weil im Kindergarten und am Ende der Messreihe der prognostizierte und der tatsächlich erreichte Durchschnittswert in den Pilotierungen nicht signifikant unterschiedlich waren.

Zusammenfassend kann gesagt werden, dass für die Leistungserfassung drei adaptive Aufgaben eingesetzt werden, die in jedem Jahr einen vergleichbaren Schwierigkeitsgrad haben, so wie dies im täglichen Schulbetrieb auch der Fall ist: Um eine gute Note zu erzielen muss man absolut gesehen in der zweiten Klasse mehr leisten als in der ersten. Nicht in Erwägung gezogen wurden als Messung zur Erfassung der Leistungsfähigkeit Tests, die die Lehrkräfte selber erstellen. Im Kindergarten wird das ja gar nicht gemacht und auch im ers-

ten Schuljahr, wenn die Kinder die nötige Schreibfertigkeit erst erwerben, ist das sehr selten. Es ist aber auch aus pädagogischen Gründen für Lehrkräfte der Schuleingangsstufe schwierig, sehr selektive Leistungstests durchzuführen, von externen „Weisskitteln“ wird dies im schulischen Umfeld besser akzeptiert. Im Gegensatz zu Leistungsmessungen, die eine Lehrperson durchführt, waren die ersten Aufgaben auf einem Niveau, das deutlich unterhalb schwächerer Leistungen der aktuellen Klassenstufe lag. Dann wurden die Aufgaben schwieriger und die Befragung endete erst mit dem Misserfolg des Kindes. Leistungen, wie sie für höhere Schuljahre erwartet wurden, konnten so erreicht werden. Die Leistungsmessung hörte also nicht bei den besten Leistungen der aktuellen Klassenstufe auf.

Die zwei Sportaufgaben wurden aus einem bestehenden Testverfahren übernommen, dies auch auf Empfehlung Stefan Valkanover's (mündliche Mitteilung) der auf Grund seiner Experimente zum Schluss kam, dass das „Rückwärts balancieren“ und das „monopedale Überhüpfen“ für die Leistungsfähigkeit im Sport sehr repräsentativ sind und bei denen Kinder Leistungsunterschiede innerhalb der Gleichaltrigengruppe gut wahrnehmen können. Eine der beiden Aufgaben musste aus zeitlichen Gründen verändert werden um die Durchführung pro Klasse auf einen halben Tag beschränken zu können. Diese Vorgabe wollten wir einhalten, um den informellen Austausch der Ergebnisse unter den Kindern möglichst in jedem Jahr konstant und generell niedrig zu halten. Den Kindern wurde deshalb abverlangt, dass sie gegenüber ihren Mitschülern bis zur Mittagspause ihre Leistungen und die Einschätzung nicht mitteilen. Die Kinder wurden über den gesamten Untersuchungszeitraum optisch überwacht um diese Regel auch durchzusetzen. Über die Mittagspause, auf dem Schulweg beispielsweise, hätte das aber nicht kontrolliert werden können. Pro Klasse wurden darum 4 – 5 Untersuchungspersonen eingesetzt, damit die Durchführung in einem Halbtage möglich war. Die Sportaufgaben wurden aus zeitlichen und häufig auch aus räumlichen Notwendigkeiten jeweils im Flur oder einem freien Schulzimmer durchgeführt. Der Versuchsleitende holte jeweils ein Kind im Klassenzimmer persönlich ab, führte es durch alle drei Aufgaben und am Schluss wieder ins Klassenzimmer zurück. Die Studierenden, die ja alle in der Ausbildung zur Lehrperson standen hatten bereits mehrwöchige Erfahrungen mit Schulkindern, sodass es ihnen durchwegs gelang, mit den Kindern einen angemessenen Kontakt aufzubauen, ihre Motivation zu unterstützen und ihre Ergebnisse nicht wertend zu kommentieren. Die Verwendung dreier für den deutschsprachigen Raum genormten Messverfahren erlaubt eine Replizierung auch in anderen Ländern.

3.4.3 Die Operationalisierung der Selbsteinschätzung

Die noch eingeschränkten kognitiven Fähigkeiten und Fertigkeiten von Kindern im Schuleingangsalter bedingen klar differenzierende Fragen, die ein hohes Mass an Anschaulichkeit besitzen. In einem Überblicksartikel analysieren und vergleichen Asendorpf und van Aken (1993) verschiedene Messinstrumente zur Erfassung der Selbsteinschätzung. Für junge Kinder im Vorschulalter ist es wichtig, dass die Skala einen visuellen Bezug zulässt und möglichst konkret ist. Auf der Kindergartenstufe wurden deshalb bereits in der Vergangenheit anschauliche Skalen verwendet. Helmke (und andere) benutzten beispielsweise 5 kleine Holzfiguren, die Kinder darstellten und die eine Leistungsrangreihe der eigenen Schulklasse repräsentierten. Das Kind setzte dann eine andersfarbige Holzfigur neben diejenige Figur, die den eigenen Platz in der Leistungshierarchie darstellen sollte. Die oberste Figur wurde dabei als „Dies ist der Beste“, die unterste als „Dies ist der Schlechteste“ bezeichnet. In der Untersuchungspraxis von Längsschnittuntersuchungen erwiesen sich solche Skalen als valide, sofern es gelang den Iteminhalt, den Antwortmodus und das Itemsetting möglichst konstant zu halten (Helmke, 1991).

Allerdings handelt es sich dabei im Wesentlichen um anschaulich dargestellte Rangreihen, also um eine Skala, die vorwiegend zum sozialen Vergleich ermuntert. Die Fokussierung der Einschätzung an der eigenen Leistungsfähigkeit gerät so leicht in den Hintergrund, diese stellt aber die eigentliche Bezugsnorm des Schuleingangsalters dar. In der Vergangenheit wurden deshalb verschiedene Instrumente entwickelt, die mit mehreren Skalen neben dem komparativen Urteil auch das prospektive Leistungsselbstbild erfassten („Wie gut bist du, wenn du deine beste Mathematikleistung erbringen kannst?“), so beispielsweise bei Helmke (1992). In den Münchner Studien wurden aber auch die „Pictorial-Scale“ von Harter und Pike (1984) und das „Self-Perception-Profile“ (Harter 1985) in deutschen Versionen eingesetzt. Asendorpf und van Aken (1995) kommen in ihrer Studie zum Schluss, dass diese Skalen ab zweitem Schuljahr eingesetzt werden können, um das Selbstkonzept in spezifischen Bereichen zu erfassen. Für jüngere Kinder diskutieren sie das „Self-Description-Questionnaire“ von Marsh et al. (1984). Dieses umfasst aber 64 Items und ist in der Erhebung so aufwändig, dass es zwei Sitzungen braucht. Für die vorliegende Untersuchung erweisen sich die bestehenden Skalen damit als zu zeitaufwändig und zu anspruchsvoll für die Versuchsleitenden. Eine Anlehnung an die Skala von Helmke (1991) scheint deshalb sinnvoll.

Skalen umfassen meist die Leistungsbandbreite der eigenen Klasse. Für unsere Untersuchung wollten wir es Kindern aber auch ermöglichen, sich leistungsmässig den Kindern zuzuordnen, die bereits ein Jahr älter sind, beispielsweise: „Ich gehöre sogar zu den Besten, die ein Jahr älter sind“. Mit diesem Vorgehen sollten einerseits Ceiling-Effekte vermieden

werden, andererseits gilt es der Tatsache Rechnung zu tragen, dass die Leistungen im Schuleingangsalter sehr breit streuen und es keine Seltenheit ist, dass Kindergartenkinder bereits Leistungen erbringen, die auch für Erstklässler sehr gut sind: Einige Kinder können beispielsweise schon im Kindergarten lesen und schreiben. Wie die Leistungsmessung sollte also die Einschätzskala möglichst über den eigenen Klassenrahmen hinaus gehen. Es wurde eine Treppe gezeichnet, auf der sich Kinder befanden. Auf der obersten Stufe stand das leistungsstärkste, auf der untersten das leistungsschwächste Kind (Abb.3-4-1, verkleinert; vgl. Titelbild, respektive Anhang für die Originalgrösse). In der Mitte standen die Kinder der eigenen Klassenstufe, erkennbar an den gleich eingefärbten Kleidern. Etwas verschoben oberhalb dann die Kinder, die bereits im nächsten Schuljahr sind in einer anderen Farbe. Damit einerseits ersichtlich ist, dass es Überschneidungen gibt, die Klassenstufen aber optisch gut unterschieden werden können, wurde eine Treppenskala mit 17 Stufen gebildet. Um Fehler bei der Kodierung zu vermeiden, befand sich unterhalb jeder einzelnen Treppenstufe ein Buchstaben. Die Einfärbung der einzelnen Klassen wurde auf jedem Instruktionsblatt variiert. Um die Normalverteilung abzubilden, wurden auf den mittleren Rangplätzen jedes Jahres mehrere Kinder dargestellt, auf einzelnen Stufen fehlten Kinder, um damit die Normalverteilung und die Trennung der Klassen optisch zu unterstützen. Mit dieser Treppenstufenskala wurde eine Pilotierung bei zwei Kindergartenklassen durchgeführt, um die Verständlichkeit der Instruktion und das Handling für die Versuchsleitenden zu prüfen. Anschliessend entstand zuhanden der Versuchsleitenden ein Instruktionsmanual, das die verbale Kommunikation und den genauen Ablauf der gesamten Erhebung festlegte:



Abbildung 3-4-1: Die Treppenskala (verkleinert, Originalgrösse im Anhang)

Die Versuchsleiterin zeigte dem Kind die Treppenskala und erläuterte diese in der ersten Klasse beispielsweise mit folgenden Worten:

„ Du siehst hier im Zimmer drei Aufgaben, die du zu lösen hast. In einer Aufgabe geht es darum, einen Schaumgummiblock zu überspringen, bei einer andern Aufgabe muss man über ein schmales werdendes Brett gehen und in der dritten geht's um das Rechnen.“

„Genau das siehst du hier auf dem Bild: Hier in der Mitte sind die Erstklässler, ihre Kleider haben die Farbe rot: Viele können manchmal gut und manchmal nicht so gut rechnen oder springen (zeigt auf N), andere sind recht gut (O), gut oder sehr gut (R) und sogar besser als die Zweitklässler, die hier oben stehen! Je besser, desto weiter oben, ja es geht natürlich noch weiter als bis zu den Zweitklässlern.

Andererseits gibt es Kinder, die können es noch nicht so gut oder gar nicht gut, ja die können sogar schlechter rechnen als Kinder, die noch im Kindergarten sind.

Also hier in der Mitte sind die Erstklässler, die das Rechnen manchmal gut und manchmal nicht so gut können und wenn du jetzt so schaust, wo wärst denn du auf der Treppe?“

Das Kind zeigt nun beispielsweise auf die Treppenstufe „O“. Der Versuchsleiter verifiziert die Antwort des Kindes:

„Du meinst also, dass du recht gut rechnen kannst und zwar etwas besser als die Erstklässler, die hier in der Mitte stehen?“

Wenn sich das Kind nun korrigiert, dann verifiziert der Versuchsleiter von Neuem. Nach dieser ersten Einschätzung fährt der Versuchsleiter mit der Einschätzung der nächsten Aufgabe, gemäss dem Manual, weiter (vgl. im Anhang).

Tabelle 3-4-1: Skalierung und normative Hinweise zu jeder Treppenstufe
am Beispiel der Instruktion im ersten Schuljahr

Treppenbuchstaben und Wert	Hinweis auf absolute Norm	Hinweis auf soziale Norm
E, -8	Gar nicht gut	Wie der schwächste Kindergärteler
F, -7	Gar nicht gut	Bei den schwächsten Kindergärteler
G, -6	Gar nicht gut	Schlechter als die meisten Kindergärteler
H, -5	Gar nicht gut	So wie viele Kindergärteler
I, -4	Gar nicht gut	Wie die schlechtesten Erstklässler
K, -3	Nicht so gut	Schlechter als die meisten Erstklässler
L, -2	Nicht so gut	Schlechter als viele Erstklässler
M, -1	Manchmal nicht so gut	Etwas schlechter als die mittleren Erstkl.
N, 0	Manchmal gut, manchmal nicht so gut	Hier in der Mitte der Erstklässler
O, +1	Recht gut	Etwas besser als die mittleren Erstklässler
P, +2	Gut	Besser als viele Erstklässler
Q, +3	Gut	Besser als die meisten Erstklässler
R, +4	Sehr gut	Wie die besten Erstklässler
S, +5	Sehr gut	So gut wie viele Zweitklässler
T, +6	Sehr gut	Besser als die meisten Zweitklässler
U, +7	Sehr gut	Bei den sehr guten Zweitklässler
V, +8	Sehr gut	Wie der beste Zweitklässler

Die Sozialnormhinweise wurden für jedes Untersuchungsjahr der Situation entsprechend adaptiert: Im ersten Jahr standen die Kindergartenkinder in der Mitte, im folgenden Jahr Schülerinnen und Schüler des ersten Schuljahres, usw. Das Durchführungsmanual umfasste gleichzeitig das Datenerhebungsblatt, die Antworten der Kinder (Treppenstufe) wurden mit einem Buchstaben codifiziert, der dann zur Eingabe in den entsprechenden Zahlenwert umgerechnet wurde (beispielsweise: „U“ als +7).

Aus den bisherigen Befunden leiteten wir ab, dass die eigene Leistungsfähigkeit im Schuleingangsalter im Wesentlichen vom individuellen Leistungsfortschritt abgeleitet wird. Im zweiten Schuljahr gewinnt dann auch der Vergleich mit den Mitschülerinnen und -schülern an Bedeutung, während die Urteile der Lehrkraft bis zu diesem Zeitpunkt noch wenig in die Selbsteinschätzung einfließen. In der Instruktion der Treppenskala gilt deshalb der absoluten Norm die verbale Aufmerksamkeit, der Vergleich zur sozialen Gruppe ist vor allem visuell sichtbar, dann aber auch jeweils im präzisierenden Nebensatz der Versuchsleitenden. Die Instruktion wurde für jedes Jahr stabil gehalten, um unkontrollierte Einflussvariablen stabil zu halten, obwohl im zweiten Schuljahr eine gleichwertige Betonung von absoluter und sozialer Norm vielleicht angepasster gewesen wäre.

Durch die gezielte Schulung und das Coaching der Versuchsleitenden, sowie die Standardisierung der Instruktion ist die Durchführungsobjektivität optimiert, die Auswertungsobjektivität ist ebenfalls gegeben. Da ähnliche Skalen bereits von anderen Autoren erfolgreich eingesetzt wurden, kann die Interpretationsobjektivität als hoch bezeichnet werden.

Da wir nur über ein Item zur Erhebung der Höhe der Selbsteinschätzung verfügen, kann die Reliabilität nicht im üblichen Sinne überprüft werden, allenfalls wäre es möglich gewesen, nach der Pilotierung im Kindergarten zwei oder drei Wochen später eine Wiederholung vorzunehmen. Bis zur ersten Erhebungswelle blieb dazu zu wenig Zeit, zudem hätten Erinnerungseffekte, die bei einjährigem Abstand nicht gleich ausgefallen wären, bei den Kindern nicht ganz ausgeschlossen werden können. Eine gewisse Stabilität der Einschätzungen von Jahr zu Jahr kann trotzdem erwartet werden, es ist beispielsweise ziemlich unwahrscheinlich, dass ein Kind im Kindergarten bei den leistungsstärksten Kindern ist und ein Jahr später dann bei den Schwächsten. Schätzt sich ein Kind nun auf der Treppenskala so extrem unterschiedlich ein, dann kann das heissen, dass das Kind tatsächlich Mühe hat, sich selbst einzuschätzen oder dass das Messinstrument wenig verlässlich ist. Es interessiert also, in welchem Mass sich die Kinder über die Zeit hinweg konsistent einschätzen, die Schätzungen also stabil sind. Dafür liefert die Korrelation der Einschätzungen über die Zeit Hinweise.

Es ist anzunehmen, dass die fehlenden Vergleichsmöglichkeiten im Fachbereich Mathematik im Kindergarten zu grösserer Schätzungenauigkeit führen als in den beiden Schuljahren, also nach einem, respektive zwei, Jahren Unterrichtserfahrung. Ein Kind kann beispielsweise feststellen, dass es zwar sehr gut ist, ein Teil der Mitschüler aber ebenfalls und dass deshalb die Einschätzung aus dem Kindergarten (z. B. „Ich bin sehr gut, ich gehöre sogar zu den besten Erstklässlern“ -> Treppenstufe 8) in der ersten Klasse nach unten korrigiert werden muss (z. B. „Ich bin sehr gut, so gut wie die besten meiner Klasse“ -> Treppenstufe 4). Das bedeutet nun, dass nur mittlere Stabilitäten erwartet werden, weil sich ja in

unseren Annahmen die Leistungserfahrung auf die Einschätzung auswirkt. Zwar gilt allgemein eine Stabilität von 0.40 bis 0.60 als mässig, allerdings unter der Voraussetzung, dass gar keine Veränderung über die Zeit erwartet wird. Stabilitäten um 0.40 können für die Korrelation der Einschätzungen über ein Jahr in unserem Falle als befriedigend betrachtet werden, weil ja Kinder auch tatsächlich in Bezug auf ihre Klasse leistungsmässig auf- oder absteigen.

Als letztes Messinstrument wird eine Einschätzung der Leistungsfähigkeit an der Aufgabe erhoben. Im Balancieren rückwärts und im monopodalen Überhüpfen schätzen sich die Kinder direkt vor der Leistungserbringung ein: Wieviele Schaumgummis kann ich wohl überhüpfen? Wie weit kann ich rückwärts balancieren? Da sich in diesen Aufgaben die durchschnittliche Leistungsfähigkeit in jedem Jahr verbesserte, erhielten die Kinder einen Hinweis, einem Zweitklässler wurde beispielsweise folgendes mitgeteilt:

„Die Erstklässler springen meisten etwa so hoch (die Höhe von 5 Schaumgummis zeigen), Zweitklässler, so wie du, springen meistens etwa so hoch (6 Schaumgummis zeigen, Drittklässler meistens etwa so hoch (8 zeigen), aber es gibt ganz selten auch einen Erstklässler der so hoch springt (11 zeigen). Was glaubst du, wie hoch kannst du springen?“

3.5 Der Untersuchungsplan

Die folgende Tabelle zeigt, zu welchen Zeitpunkten welche Teile der Stichprobe erhoben wurden. Aufgeführt sind ebenfalls Pilotierungen und Kontrollgruppenerhebungen.

Tabelle 3-4-2 : Untersuchungsplan

Wann	Was	Personen
März 1999	Pilotierung erste Erhebung	2 Klassen
Juni 1999	Erste Erhebung	Längsschnittgruppe, 14 Kl., 221 Vpn
November 1999	Pilotierung „rückwärts Balancieren“	2 Erste Klassen
Mai 2000	Zweite Erhebung	Kontrollgruppe 1, 101 Vpn
Juni 2000	Dritte Erhebung	Kontrollgruppe 2, 43 Vpn
Mai / Juni 2000	Zweite Erhebung	Längsschnittgruppe, 14 Kl., 210 Vpn
Juni 2001	Dritte Erhebung	Kontrollgruppe 2, 53 Vpn
Mai / Juni 2001	Dritte Erhebung	Längsschnittgruppe, 14 Kl., 200 Vpn

Die Daten von 43 Kindern der Kontrollgruppe 2 wurden bereits 2000 erhoben, um damit eine Pilotierung der dritten Erhebung zu gewinnen. Die Befunde führten dazu, dass die dritte Erhebung unverändert durchgeführt wurde. In jedem Jahr gibt es 5 – 10 Kinder, die am vereinbarten Untersuchungstag krank oder sonst abwesend waren. Diese Kinder wurden innert 3 Wochen nachuntersucht, um ihren Datensatz zu erhalten. Bei längerer Abwesenheit, viel das Kind aus der Untersuchungsreihe.

4. Darstellung der Befunde und Interpretation

4.1. Übersicht

Die Darstellung der Befunde erfolgt in vier Unterkapiteln, beginnend mit den Befunden zur Mathematik, gefolgt von denjenigen zum monopedalen Überhüpfen und dem Balancieren rückwärts und einer abschliessenden Gesamtschau. Im folgenden Kapitel zur Mathematik werden zuerst die Befunde zur Einschätzung der eigenen Leistungsfähigkeit im Vergleich zu den anderen Kindern (soziale Bezugsnorm) dargestellt. Daran anschliessend werden die tatsächlichen Leistungen der Kinder präsentiert. In einem weiteren Schritt werden die Leistungen der Kinder zu deren Einschätzung in Bezug gebracht, woraus dann Aussagen über Überschätzungs-, respektive Unterschätzungstendenzen der Kinder resultieren. Schliesslich erfolgt noch die Prüfung der Korrelation von Leistung und Einschätzung und deren Veränderung über die drei Messzeitpunkte hinweg, also die Entwicklung der Schätzgenauigkeit. Das letzte Kapitel präsentiert die Ergebnisse aller drei Untersuchungen zusammenfassend und schliesst mit einer Gesamtdiskussion und einem Ausblick.

4.2 Ergebnisse in Mathematik

Zuerst erfolgt die Darstellung der Mittelwerte der Einschätzhöhe und der Verteilungen der Häufigkeiten. Mit Signifikanzprüfungen gilt es festzustellen, ob die Hypothese 1.1., wonach sich Kinder in diesem Alter signifikant über dem prognostizierten Durchschnitt einschätzen, bestätigt werden kann. Anschliessend wird untersucht, ob sich die Mittelwerte zwischen den verschiedenen Jahren signifikant unterscheiden und ob sich Milieu- und Geschlechtsunterschiede ergeben. Dies erfolgt mit zwei verschiedenen Verfahren, aber in jeweils einem einzigen Schritt: Zuerst mit einer ANOVA mit der Einschätzung der Leistung als abhängiger Variablen und dem Geschlecht, der Klassenstufe sowie dem Milieu als unabhängigen Variablen. Mittels einer UNIANCOVA wird die gleiche Berechnung dann nochmals vorgenommen, aber klassenstufenweise und mit der Leistung als Kovariaten. Beide Male handelt es sich um einen mehrfaktoriellen Plan mit Messwiederholung der abhängigen Variablen, wie auch der Kovariaten. Das erste dieser beiden Verfahren liefert Befunde darüber, wie sich die

Kinder gesamthaft einschätzen und ob sich klassenstufenweise Unterschiede ergeben. Das zweite Verfahren (UNIANCOVA) partialisiert den Einfluss der Leistung aus, die ja, gemäss unseren Hypothesen, vermutlich signifikant mit der Einschätzung korreliert. Allerdings sind jetzt nur noch klassenstufenweise Aussagen möglich. Allfällige Leistungsunterschiede zwischen den Geschlechtern oder dem Milieu können aber nun kontrolliert werden.

Neben der Frage, wie hoch sich Kinder während der Schuleingangsphase einschätzen, interessiert auch, ob diese Einschätzung so hoch ist, dass es sich dabei um eine Überschätzung (es könnte auch eine Unterschätzung sein) handelt. Um das zu beantworten, wurde von der Einschätzung die Leistung direkt in Abzug gebracht und der verbliebene Rest (als negative oder positive Überschätzung) betrachtet.

Ein weiteres Kapitel untersucht die Korrelationen zwischen Leistung und Einschätzung der Leistung, denn Einschätzungshöhe und Schätzgenauigkeit sind logisch und statistisch voneinander unabhängig, die Schätzgenauigkeit wird deshalb in einem eigenen Unterkapitel untersucht.

Bei der Argumentation in den verschiedenen Kapiteln kann es manchmal von Nutzen sein, sich über die Verteilung der Häufigkeiten genauer informieren zu können. Das ist im letzten Kapitel dargestellt. Dort finden sich auch Einzelbefunde, beispielsweise zum „Schicksal“ derjenigen Kinder, die sich als sehr schwach einschätzen und die über ein entsprechend negatives Leistungsselbstbild in Mathematik verfügen.

4.2.1 Höhe der Einschätzung im Kindergarten und den beiden Schuljahren

Tab.4-2-1 zeigt die Befunde für die 200 Kinder, die vollständig an der Untersuchung teilgenommen haben. Die Kinder konnten beispielsweise antworten, dass sie sich „manchmal gut, manchmal nicht so gut, ungefähr in der Mitte meiner Klassenstufe“ beurteilen. Diese Aussage erhält den Wert 0 und entspricht genau der erwarteten Leistungsmitte des Altersjahrgangs.

Tabelle 4-2-1: Lageparameter und Streuungsmasse der Einschätzung
in Mathematik im Kindergarten, der ersten und zweiten Klasse

	Schätzung der eigenen Leistung, Kindergarten	Schätzung der eigenen Leistung, 1.Klasse	Schätzung der eigenen Leistung, 2.Klasse
N	200	200	200
Arithm. Mittel	0.50	1.62	1.16
Median	0	1	1
Modus	0	0	0
Std. Abweichung	2.925	2.539	2.373
Varianz	8.553	6.448	5.633
Minimum	-7	-4	-8
Maximum	8	8	8

(Skala: -8 bis +8; die Werte meinen jeweils eine Treppenstufe)

Im Durchschnitt haben sich die Kinder in allen Jahren über dem erwarteten Skalenmittelwert eingeschätzt, die arithmetischen Mittel der Einschätzungen liegen nämlich bei 0.50, 1.62 und 1.16. Alle diese Ueberschätzungen sind signifikant, die t-Tests ergeben für den Kindergarten $t(199) = 2.394$, $p < 0.05$, für die erste Klasse $t(199) = 9.022$, $p < 0.01$ und für die zweite Klasse $t(199) = 6.912$, $p < 0.01$.

Die Mediane befinden sich im Kindergarten in der Skalenmitte, in der ersten und zweiten Klasse aber eine Stufe darüber. Im Kindergarten sind es 70 Kinder (35%), die sich unterhalb des Medians sehen, 53 Kinder (26.5%) sehen sich genau in der Mitte und 77 (38.5%) Kinder liegen über der Mitte. In der ersten und zweiten Klasse liegt der Median eine Stufe höher: Nur noch 89 Kinder (44,5%; erstes Schuljahr), respektive 98 Kinder (49% im 2. Schuljahr) schätzen sich in der Skalenmitte (=0) oder tiefer ein. Die Einschätzungen verschieben sich also deutlich, in den beiden Schuljahren schätzen sich zunehmend mehr Kinder oberhalb der Skalenmitte ein.

Der Modus ist durchwegs 0, die grösste Anzahl Einzeleinschätzungen der Leistung findet sich also genau in der Skalenmitte. Die Zahl der Kinder, deren Angaben auf den Modus fällt, steigt von Klassenstufe zu Klassenstufe, von 53 über 58 auf 63 (Abbildung 4-2-1).

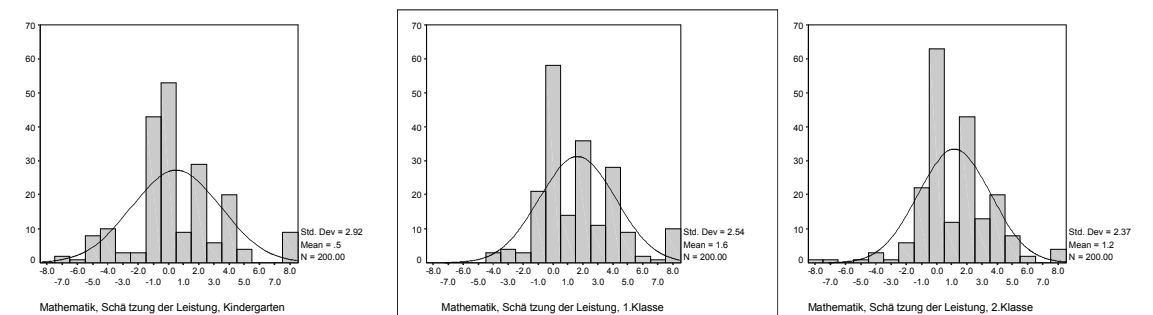


Abbildung 4-2-1: Einschätzung der eigenen Leistungsfähigkeit in Mathematik
vom Kindergarten zur zweiten Klasse

Die Varianz der Kindergartenkinder unterscheidet sich signifikant von den beiden Schuljahren (KG-1.Kl.: $F(1.26)=1.326$, $p<0.05$; KG-2.Kl.: $F(1.39)=1.518$, $p<0.01$), der Unterschied zwischen erstem und zweitem Schuljahr ist nicht signifikant ($F(1.20)=1.145$, $p>0.10$). Im Kindergarten streuen also die Selbsteinschätzungen signifikant stärker als später, was aufgrund der mangelnden Unterrichtserfahrung in diesem Fach zu erwarten war. Da die Stichprobengröße hoch ist, die Anzahl der untersuchten Kinder identisch, und die Verteilung symmetrisch, kann der Einfluss der Varianzunterschiede auf die ANOVA-Berechnungen vernachlässigt werden (Bortz, 1999, S.263).

Die Hypothese 1.1., wonach sich die Kinder aller Klassenstufen in Mathematik im Durchschnitt selbst signifikant höher einschätzen als sie den Durchschnitt ihrer Klasse einschätzen, wird damit bestätigt.

4.2.2 Befunde zur Einschätzhöhe unter Berücksichtigung der Variablen Klassenstufe, Geschlecht und Milieu

Mittels einer dreifaktoriellen ANOVA wurden die Befunde zu den Hypothesen 2, 4 und 5 in einem einzigen Schritt berechnet. Die unabhängigen Faktoren waren „Geschlecht“ (Jungen versus Mädchen, also „between subjects“) und „Milieu“ (Stadt versus Land, ebenfalls „between subjects“) und die Klassenstufe (Kindergarten versus erste Klasse versus zweite Klasse, „within subjects“).

Tabelle 4-2-2: Dreifaktorielle ANOVA zur Einschätzung der eigenen Leistungsfähigkeit in Mathematik

Quelle	df	F	p
Klassenstufe	2	9.672	0.000
Geschlecht	1	4.754	0.030
Milieu	1	0.037	0.847
Stufe x Geschlecht	2	1.096	0.335
Stufe x Milieu	2	1.974	0.140
Geschlecht x Milieu	1	0.001	0.973
Stufe x Geschlecht x Milieu	2	2.533	0.081

Keine der Interaktionen ist signifikant. Die Haupteffekte können also direkt diskutiert werden. Von diesen sind der Klassenstufen- und der Geschlechtseffekt signifikant. Die nachträgliche Kontrastanalyse zeigt, dass die Werte im Kindergarten von den Werten der

beiden anderen Klassenstufen signifikant abweichen (KG-1.Kl.: 1.12 Treppenstufen, signifikant mit $p=0.00$; KG-2.Kl.: 0.66 Treppenstufen, signifikant mit $p=0.014$), während das Absinken der Einschätzung von der ersten zur zweiten Klasse nicht signifikant ist (1.Kl.-2.Kl.: 0.46 Treppenstufen, n.s. mit $p=0.107$) (vgl. Abb.4-2-2).

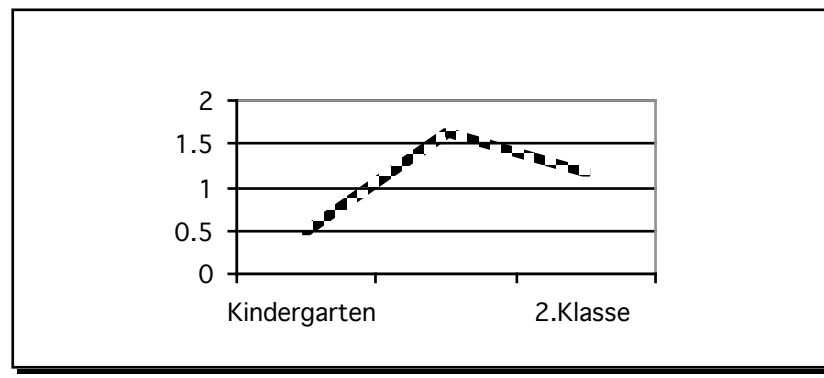


Abbildung 4-2-2: Verlauf der Einschätzungshöhe in Mathematik
(0 = „ich bin manchmal gut und manchmal nicht so gut in Mathematik,
und gehöre hier zu den mittleren meines Jahrganges“)

Wir hatten angenommen, dass die Einschätzungshöhe über die drei Klassenstufen zuerst zunimmt und dann sinkt, sodass sie in der zweiten Klasse immer noch höher ist, als im Kindergarten. Dieser Verlauf ist eingetreten, das Absinken vom ersten zum zweiten Schuljahr ist so gering, dass die Einschätzungshöhe am Ende signifikant höher ist als im Kindergarten.

Die Hypothese 2.1 ist teilweise bestätigt:

- a) Die durchschnittliche Einschätzungshöhe ist im ersten Schuljahr am höchsten
- b) sie nimmt vom Kindergarten zur ersten Klasse signifikant zu.
- c) Die Abnahme von der ersten Klasse zur zweiten Klasse ist jedoch nicht signifikant.

Die Hypothese 2.2 wird damit bestätigt: Die Einschätzungshöhe nimmt vom Kindergarten zur zweiten Klasse signifikant zu

4.2.3 Geschlechts- und Milieuunterschiede

Insgesamt ergibt sich in der Einschätzhöhe ein signifikanter Geschlechtsunterschied ohne Interaktionen ($p=0.03$, vgl. Tab.4-2-2). Die Knaben schätzten sich durchschnittlich höher ein als die Mädchen, wie Tab.4-2-3 und Abb.4-2-3 zeigen:

Tabelle 4-2-3: Mittelwerte der Einschätzung nach Klassenstufen, Geschlecht und Milieu

	Insgesamt (N=200)	Mädchen (N=98)	Knaben (N=102)	Stadt (N=94)	Land (N=106)
Insgesamt Aller Jahre	1.090	0.898	1.281	0.996	1.183
Kindergarten	0.49	0.35	0.64	0.74	0.27
1.Klasse	1.62	1.23	1.99	1.46	1.76
2.Klasse	1.16	0.84	1.47	1.14	1.18

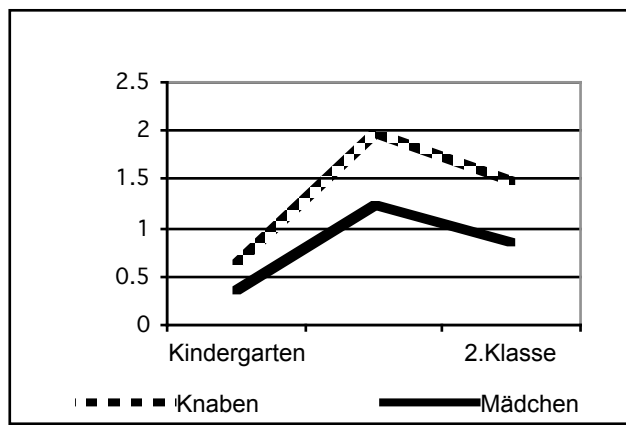


Abbildung 4-2-3: Verlauf der Einschätzhöhe in Mathematik

Die Hypothese 4.1., wonach sich Knaben in Mathematik leistungsmässig höher einschätzen als Mädchen kann bestätigt werden

Der Milieuunterschied ist nicht signifikant ($F=0.037$, $p=0.847$) womit sich Einzelvergleiche erübrigen. Weil in der vorliegenden Untersuchung diesem Aspekt besondere Aufmerksamkeit geschenkt wird und weil entsprechende Hypothesen formuliert sind, werden die Ergebnisse hier noch kurz kommentiert und dargestellt. Im Kindergarten sind, wie erwartet die Einschätzungen der Landkinder tiefer als diejenigen der Stadtkinder (aber nicht signifikant),

nachher ist es umgekehrt und insgesamt werden die Unterschiede kleiner, sie sind ebenfalls nicht signifikant (Abb.4-2-4).

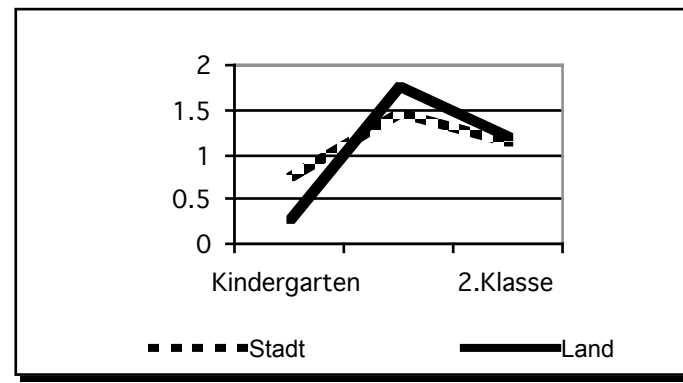


Abbildung 4-2-4: Durchschnittliche Abweichung der Einschätzung von Stadt- und Landkindern (1= eine Treppenstufe über dem Klassenmittel)

Hypothese 5.1, wonach urbane Schulklassen eine signifikant höhere Einschätzung haben als rurale kann nicht bestätigt werden.

4.2.4 Mittelwerte der Einschätzhöhe unter Berücksichtigung der Leistung und der unabhängigen Variablen Geschlecht und Milieu

Knaben und Mädchen unterscheiden sich in der Einschätzhöhe signifikant. Das heisst allerdings noch nicht, dass sich die Knaben gegenüber den Mädchen signifikant stärker überschätzen. Es könnte ja sein, dass auch ihre Leistungen höher sind. Um einen solchen allfälligen Unterschied zu kompensieren wurde die Leistung (Rohwerte im KAB) kontrolliert. Durch je eine pro Klassenstufe durchgeführte UNIANCOVA wurden neben dem Geschlecht auch die unabhängige Variable Milieu, sowie allfällige Interaktionen überprüft (Tab.4-2-6).

Tabelle 4-2-4: Klassenstufenweise UNIANCOVA mit Geschlecht und Milieu als unabhängigen Faktoren und der Leistung als Kovariaten

Quelle	Df	F	P
Kindergarten			
Leistung KG	1	17.405	0.001
Geschlecht	1	0.095	0.758
Milieu	1	0.373	0.542
Geschlecht x Milieu	1	2.689	0.103
Erste Klasse			
Leistung erste Klasse	1	5.376	0.021
Geschlecht	1	3.776	0.053
Milieu	1	1.884	0.172
Geschlecht x Milieu	1	0.108	0.743
Zweite Klasse			
Leistung zweite Klasse	1	24.491	0.000
Geschlecht	1	1.299	0.256
Milieu	1	0.611	0.435
Geschlecht x Milieu	1	1.359	0.245

Die Interaktionen sind nicht signifikant. Die Haupteffekte sind darum direkt diskutierbar. In jedem Jahr hat die Leistung einen signifikanten Varianzanteil an der Einschätzung. Das deutet auf signifikante Korrelationen zwischen Leistung und Einschätzung hin (vgl. weiter unten). Geschlecht und Milieu haben, unter Kontrolle der Leistung, keinen signifikanten Effekt auf die Einschätzungshöhe. Eine Betrachtung der einzelnen Ausbildungsjahre (vgl.: Tab.4-2-4) ergibt nur in der ersten Klasse einen beinahe signifikanten Geschlechtsunterschied ($p=0.053$).

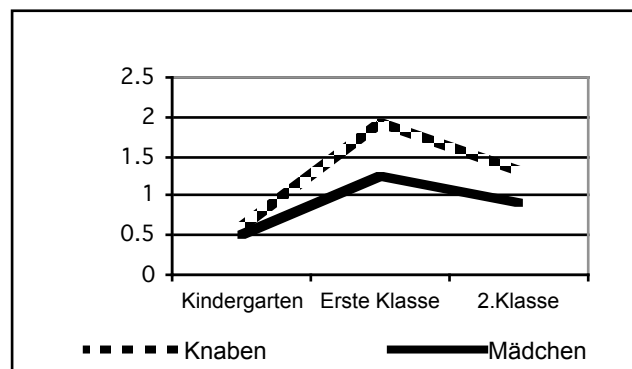


Abbildung 4-2-5: Durchschnittliche Abweichung der Einschätzung (unter Kovarianz der Leistung) von Knaben und Mädchen (1= eine Treppenstufe über dem Klassenmittel)

Wenn also die Leistung neben der Einschätzung kontrolliert werden kann (weil sie gemessen wurde), dann ergibt sich in der vorliegenden Untersuchung ein anderer Befund, der Geschlechtsunterschied in der Höhe der Einschätzung der eigenen Leistungsfähigkeit in Mathematik ist dann nicht signifikant.

Die Hypothese 4.3., wonach sich Knaben, gemessen an ihren Leistungen, höher einschätzen als Mädchen kann nicht bestätigt werden.

Unterschiede in der Einschätzhöhe von Stadt-, respektive Landkindern sind auch unter Kovariation der Leistung nicht signifikant. Weil die Befunde interessieren, werden sie trotzdem kurz diskutiert. Im Kindergarten sind die Einschätzungen der Landkinder (nicht signifikant) tiefer als diejenigen der Stadtkinder, so wie es erwartet wurde. In den beiden Schuljahren ist es dann umgekehrt: Stadtkinder schätzen sich weniger hoch ein als Landkinder, aber eben statistisch nicht signifikant (Abb.4-2-6).

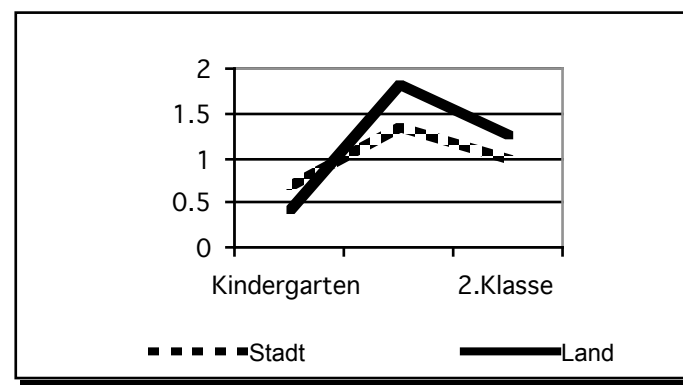


Abbildung 4-2-6: Durchschnittliche Abweichung der Einschätzung von Stadt- und Landkindern (1= eine Treppenstufe über dem Klassenmittel)

Die Hypothese 5.2., wonach sich urbane Schulklassen, gemessen an ihren Leistungen, höher einschätzen als rurale kann nicht bestätigt werden.

4.2.5 Die Leistungen in Mathematik

Unter Berücksichtigung der Leistung ergeben sich in den obigen Befunden keine signifikanten Geschlechtsunterschiede mehr. Offensichtlich sind die Leistungen von Knaben und Mädchen unterschiedlich, obwohl dies vom Messinstrument, dem KAB-C, nicht vorhergesagt

wurde. Mit einer ANOVA mit Klassenstufe, Milieu und Geschlecht als unabhängigen Variablen wurden die Leistungsunterschiede einer Signifikanzprüfung unterzogen (Tab.4-2-5).

Tabelle 4-2-5: ANOVA: Leistung in Mathematik mit Klassenstufe und Geschlecht

	Df	F	P
Klassenstufe	2	915.67	0.001
Geschlecht	1	3.859	0.051
Klassenstufe x Geschlecht	2	1.144	0.310

Die Interaktion ist nicht signifikant. Die Klassenstufen unterscheiden sich signifikant. Die nachträgliche Kontrastanalyse zeigt, dass die Leistungen in der ersten Klasse signifikant (6.5 Rohwertpunkte, sig. mit $p=0.00$) höher sind, als im Kindergarten und in der zweiten Klasse signifikant höher als in der ersten Klasse (2.5 Rohwertpunkte, sig. mit $P=0.00$) Der Geschlechtsunterschied ist ganz knapp nicht signifikant, vermutlich, weil die Leistungsunterschiede im Kindergarten noch sehr klein sind, sie nehmen dann aber zu. Da der Geschlechtseffekt nicht signifikant ist, erübrigt sich die Analyse der einzelnen Jahre. Wie sich die Leistungsunterschiede zwischen Knaben und Mädchen verändern, ist aber von Interesse, die entsprechenden Mittelwerte finden sich in Tab. 4-2-6.

Tabelle 4-2-6: Klassenstufenweise Mittelwerte der Leistung von Knaben und Mädchen in Mathematik (in Rohwertpunkten des KAB)

	Mädchen	Knaben	Differenz In Punkten	Höhere Leistung der Knaben in %
KG	17.94	18.38	0.44	2.45 %
1.Kl.	24.27	25.04	0.77	3.17 %
2.Kl.	26.64	27.75	1.11	4.16 %

In der durchschnittlichen Leistungsfähigkeit unterscheiden sich die Geschlechter also zunehmend stärker (aber eben nicht signifikant). Würden nur die beiden Klassenstufen, in denen Mathematik unterrichtet wird, berücksichtigt, dann wäre der Geschlechtsunterschied signifikant, die Knaben wären signifikant besser als die Mädchen. Kontrolliert man in der Analyse der Leistung neben dem Geschlecht auch das Milieu ergibt sich folgendes Bild:

Tabelle 4-2-7: ANOVA: Leistung in Mathematik; Klassenstufe, Geschlecht, Milieu

Quelle	df	F	P
Klassenstufe	2	889.320	0.000
Geschlecht	1	3.066	0.082
Milieu	1	1.895	0.170
Stufe x Geschlecht	2	1.573	0.212
Stufe x Milieu	2	2.507	0.094
Geschlecht x Milieu	1	0.353	0.553
Stufe x Geschlecht x Milieu	2	0.469	0.586

Die Interaktionen sind nicht signifikant. Von den Haupteffekten ist die Klassenstufe signifikant, die Leistungen in den verschiedenen Jahren unterscheiden sich also signifikant. Alle Kontraste sind signifikant, auch der kleinste Leistungsunterschied (1.Kl.-2.Kl.:2.5 Rohwertpunkte) mit $p=0.00$. Von Jahr zu Jahr nimmt die Leistung also signifikant zu. Ein Trend ($p<0.10$) ergab sich beim Geschlecht, die Knaben waren insgesamt besser ($F(3.89)=3.066$, $p=0.082$). Der Haupteffekt „Milieu“ ist nicht signifikant, die ländlichen Kinder sind im Kindergarten schwächer als die städtischen (1.07 Rohwertpunkte, n.s.mit $p=0.094$). In der ersten Klasse nimmt der Unterschied auf 0.38 Punkte ab, in der zweiten Klasse beträgt er 0.46 Punkte. Die ländlichen Kinder sind ab Ende des ersten Schuljahres kaum leistungsschwächer als diejenigen in den städtischen Klassen. Nun folgt die Darstellung und Beschreibung der Leistungen in Mathematik für die Gesamtstichprobe (Tab.4-2-8).

Tabelle 4-2-8: Mittelwerte und Streuungsmasse der Leistungen in Mathematik (Rohwertpunkte nach KAB-C)

	Leistung, Kindergarten	Leistung, 1.Klasse	Leistung, 2.Klasse
N	200	200	200
Arith.Mittelwert	18.17	24.66	27.21
Median	18	24	27
Modus	20	24	25
Std. Abweichung	4.416	2.348	2.834
Varianz	19.505	5.512	8.033
Range	19	15	14

Die einzelnen Lageparameter zeigen, dass Median und arithmetischer Mittelwert praktisch zusammenfallen. Die Standardabweichung, respektive die Varianz ist im Kindergarten deutlich grösser als später. Die beiden Schuljahre unterscheiden sich diesbezüglich nicht signifikant ($F(1.26)=1.14$, $p>0.05$). Die Varianzen zwischen Kindergarten und erster, respektive zweiter Klasse unterscheiden sich signifikant (KG -1.Kl: $F(1.26)=1.33$, $p<0.05$ und KG-

2.Kl.: $F(1.39)=1.518$, $p<0.01$). Aus den bereits oben erwähnten Gründen können diese Unterschiede vernachlässigt werden (Bortz, 1999).

Die Leistungsdichte ist im Kindergarten, also bevor der Unterricht in Mathematik einsetzt, im Vergleich zu den beiden Schuljahren kleiner, die Leistungsunterschiede sind grösser (Range: 19 Punkte). Die vergleichsweise höhere Leistungsdichte in den beiden Schuljahren macht es für die Kinder vielleicht schwieriger, untereinander Leistungsunterschiede festzustellen, weil viele Mitschüler eine ähnliche Leistung erbringen (Abb.4-2-7). Die Range ist in den Schuljahren kleiner als im Kindergarten, weil insbesondere die Kinder, die unter dem Leistungsniveau ihres Alterjahrganges liegen, fast vollständig wegfallen. Aus schulischer Sicht interessant ist die Zunahme der Leistungsstreuung vom ersten zum zweiten Schuljahr im „Mittelfeld“ (innerhalb +/- einer Standardabweichung).

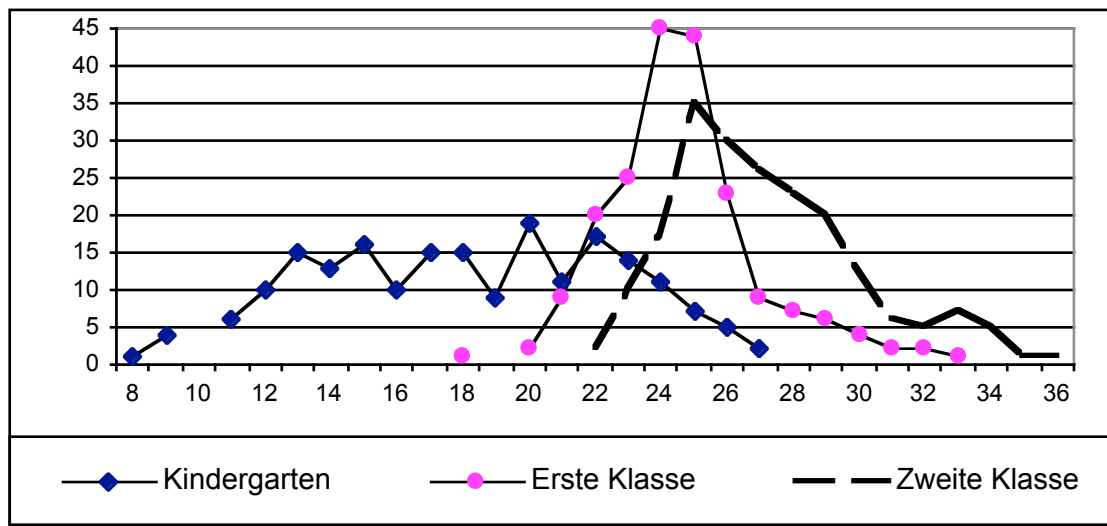


Abbildung.4-2-7: Anzahl Kinder (0-45), die eine bestimmte Leistungspunktzahl (8-36) im KAB erreicht haben, vom Kindergarten zur zweiten Klasse.

Der durchschnittliche Leistungszuwachs in Rohwertpunkten ist in der vorliegenden Untersuchung vom Kindergarten zur ersten Klasse doppelt so hoch (6.49 Punkte), wie von der ersten zur zweiten Klasse (2.54 Punkte, vgl. Tab.4-2-8).

Es stellt sich nun die Frage, ob eine so grosse Leistungszunahme den Erwartungen des KAB entspricht oder ob die Experimentalgruppe gegenüber der Normierung signifikant abweichende Leistungen erbrachte (Tab.4-2-9).

Tabelle 4-2-9: Mittelwerte der Leistungen im Vergleich zu den standardisierten Normen des KAB, jeweils in Rohwertpunkten.
Abweichung der Leistung von der KAB-Norm (= Differenz).

Leistung Nach Alter	4 Jahre 8 Monate	5 Jahre 8 Monate	Kindergarten (6J;8Mte)	1.Klasse (7J;8Mte)	2.Klasse (8J;8Mte)	3.Klasse (9J;8Mte)
KAB	9.66	13.25	18.50	23.20	27.20	29.20
Leistung			18.17	24.66	27.20	
Differenz			- 0.33	+ 1.46	0	

Im Vergleich zur KAB-Normierung sind die untersuchten Kinder im Kindergarten leistungsmässig durchschnittlich etwas schwächer, allerdings nicht signifikant, wie die Überprüfung im t-Test ergab ($t_{(199)} = -1.073$, $p > 0.10$). In der ersten Klasse dagegen sind die untersuchten Schülerinnen und Schüler hochsignifikant besser als von den KAB-Normen her zu erwarten wäre ($t_{(199)} = 8.795$, $p < 0.01$), während in der zweiten Klasse die Durchschnittswerte numerisch identisch sind.

Die Mathematikleistungen im Kindergarten und der zweiten Klasse, also am Beginn und Ende der Längsschnittuntersuchung, weichen von denen der Eichstichprobe nicht signifikant ab. Der Leistungszuwachs in KAB-C-Rohwertpunkten ist in der vorliegenden Untersuchung vom Kindergarten zur ersten Klasse sehr gross, weil die Kinder im Kindergarten ja leicht unter, in der ersten Klasse dann aber signifikant über den erwarteten Werten lagen. Die Leistungszunahme beträgt zuerst 6.49 Punkte, dann noch 2.54 Punkte. Auch gemäss der KAB-C-Normierung ist der Leistungszuwachs bei Kindern über die Jahre nicht linear (Tab.4-2-9), er nimmt beispielsweise vom Kindergarten zur zweiten Klasse jedes Jahr etwas ab. Sowohl unsere Befunde, wie auch die KAB-C-Normen deuten darauf hin, dass der Leistungszuwachs in Bezug auf mathematische Fähigkeiten im ersten Ausbildungsjahr am grössten ist. In der Schweiz beginnt der Mathematikunterricht später als in Deutschland. Das könnte die Zuwachsunterschiede zwischen der KAB-C-Normierung und der vorliegenden Untersuchung erklären, denn gemäss den KAB-C-Normen wäre der grösste Leistungszuwachs ein Jahr vorher zu erwarten (vgl.Tab.4-2-9): Die Normierungsstichprobe des KAB-C für das Alter 6 Jahre, 8 Monate, die der Kindergartenstichprobe unserer Untersuchung entspricht, besteht aus 227 (72%) Kindern des Kindergartens und 88 (28%) Kindern, die die Schule besuchen. Durch die frühere Einschulung eines ansehnlichen Teils der Kinder der Normstichprobe lässt sich die leicht höhere Leistung der Normstichprobe plausibel erklären. In unserer Untersuchung ist dementsprechend der grössere der beiden Leistungszunahmen vom Kindergarten zur ersten Klasse erfolgt. Das bestärkt die Vermutung aus Kapitel 2.3.3., dass Kinder in diesem Alter die Einschätzung der eigenen Leistungsfähigkeit am erlebten Leistungsfortschritt

messen: In unseren Befunden war die Einschätzung der Leistungsfähigkeit am Ende des ersten Schuljahres höher als ein Jahr später, so wie auch der Leistungszuwachs im ersten Jahr viel grösser war, als in der zweiten Klasse.

4.2.6 Über- oder Unterschätzung als Differenz zwischen Einschätzung und Leistung in Treppenstufen

Die bis hier diskutierte Einschätzung der eigenen Leistung bezog sich auf den Vergleich mit den anderen Schülern, respektive der vermuteten eigenen Leistungsfähigkeit. Die Leistung, eingebracht als Kovariate, wurde kontrolliert, um nicht voreilige Schlüsse über Unterschiede zwischen den Geschlechtern oder den Milieus zu ziehen. Im Folgenden wird von der Einschätzung die Leistung direkt in Abzug gebracht, sodass bei einer Differenz von einer Über-, respektive Unterschätzung gesprochen werden kann. Dieses Vorgehen ist auch zweckmässig, weil durch eine Kontrastanalyse überprüft werden soll, ob die Veränderungen in der Über-, respektive Unterschätzung zwischen den Klassenstufen signifikant abweichen. Um die Über- / Unterschätzung zu bestimmen, wurde die Leistung in Rohwertpunkten in die gemäss KAB-C-Tabelle zugehörige Stufe auf der Einschätztreppe umgerechnet. Nur bei insgesamt 11 von 600 Leistungsrohwerten mussten bei diesem Vorgehen zwei Werte zu einem Treppenwert zusammengefasst werden. Beispielsweise erhielten die Leistungsrohwerte 17 und 18 im ersten Schuljahr gemeinsam den Wert -5 aus der Treppenskala zugeordnet. In allen anderen Fällen konnte jeder Rohwert exakt einem Treppenskalawert zugeordnet werden (vgl. Tab.4-2-17). Die Überführung der Rohwerte in Treppenskalenwerte führte also zu äusserst geringem Informationsverlust und zu vernachlässigbar tiefen Verzerrungen. Die Über-, respektive Unterschätzung konnte nun berechnet werden indem Leistung, codiert als Treppenwert, von der Selbsteinschätzung auf der Treppe in Abzug gebracht wurde. Dazu ein Beispiel: Ein Mädchen der ersten Klasse schätzt sich mit dem Wert 0 ein, „in Mathematik bin ich manchmal gut, manchmal nicht so gut, wie die Erstklässler hier in der Mitte“. Im K-ABC-Test ergibt die Umrechnung der Rohwertpunkte der Leistung dieses Mädchens in eine Treppenstufe den Wert -2 („nicht so gut, schlechter als viele Erstklässler“). Dieses Kind hat sich also um zwei Stufen verschätzt. Seine tatsächliche Leistung ist tiefer als die Einschätzung, es hat sich überschätzt.

Tabelle 4-2-10: Dreifaktorielle ANOVA mit Über-/ Unterschätzung als abhängiger Variable

Quelle	df	F	P
Klassenstufe	2	4.619	0.012
Geschlecht	1	0.067	0.795
Milieu	1	2.525	0.114
Geschlecht x Milieu	1	0.452	0.502
Stufe x Geschlecht	2	0.339	0.701
Stufe x Milieu	2	0.076	0.919
Stufe x Geschlecht x Milieu	2	2.235	0.111

Da keine signifikanten Interaktionen bestehen, können die Haupteffekte direkt betrachtet werden. Ausser dem Haupteffekt für die Klassenstufe ergeben sich in der Tab.4-2-10 keine signifikanten Effekte. Weder Knaben und Mädchen noch Stadt- und Landkinder unterscheiden sich signifikant. Die nachträgliche Kontrastanalyse des Haupteffektes zeigt, dass der Mittelwert des Kindergartens sich nicht signifikant von demjenigen der beiden Schuljahre unterscheidet (KG.-1.Kl.: $p=0.808$; KG-2.Kl.: $p=0.245$), hingegen unterscheiden sich die Mittelwerte von erster und zweiter Klasse signifikant ($p=0.002$). Die Mittelwerte der Überschätzung nehmen also vom Kindergarten zur ersten Klasse ab (0.3 Treppenstufen), anschliessend steigt die Überschätzung wiederum um fast eine Treppenstufe (0.95) signifikant an.

In allen Jahren liegen die Mittelwerte über dem erwarteten Mittelwert von Null, die Kinder überschätzen sich also durchschnittlich. Entsprechend stellen die Befunde der folgenden Tabelle Überschätzungen dar (Tab.4-2-11).

Tabelle 4-2-11: Überschätzung der Leistung in Mathematik im Kindergarten, der ersten und der zweiten Klasse

	Überschätzung der eigenen Leistung, Kindergarten	Überschätzung der eigenen Leistung, 1.Klasse	Überschätzung der eigenen Leistung, 2.Klasse
N	200	200	200
Arithm. Mittel	0.31	0.01	0.96
Median	0	0	1
Std. Abweichung	4.473	3.040	2.980
Varianz	20.012	9.246	8.882
Range	21	18	16

Das arithmetische Mittel der Einschätzungen liegt nämlich bei 0.31, 0.01 und 0.96. Nur eine dieser Überschätzungen ist signifikant. Die Werte der t-Tests betragen $t_{(199)}=0.980$, $p>0.10$ für den Kindergarten, resp. $t_{(199)}=0.047$, $p>0.10$ für die erste Klasse. In der zweiten Klasse weicht die Überschätzung dann signifikant ab: $t_{(199)}=4.555$, $p<0.01$. Die durchschnittliche Überschätzung ist in der ersten Klasse am niedrigsten, ja praktisch nicht existent.

Hypothese 1.2. wird teilweise bestätigt: In der zweiten Klasse überschätzen sich die Kinder signifikant, in der ersten Klasse und im Kindergarten ist die Überschätzung nicht signifikant.

Vergleicht man nun die allgemeinen Befunde, dann ergeben sich Unterschiede, je nach dem ob wir damit die Einschätzhöhe oder die Überschätzung betrachten. Die Einschätzhöhe liegt in allen Jahren signifikant über dem erwarteten Mittelwert. Die Überschätzung liegt dagegen nur in der zweiten Klasse signifikant darüber. Im Kindergarten und der ersten Klasse überschätzen sich die Kinder nicht signifikant. Sowohl die Einschätzhöhe, wie die Überschätzung nehmen vom Kindergarten zur zweiten Klasse zu, allerdings nur bei der Einschätzhöhe signifikant. Unterschiedlich sind die Verläufe: Während die Einschätzhöhe zuerst signifikant zunimmt und dann wieder zurückgeht, sinkt die Überschätzung zuerst und nimmt dann wieder zu. Dies erklärt sich damit, dass die Kinder in der ersten Klasse einen signifikant über den Erwartungen (KAB-C-Norm) liegenden Leistungsdurchschnitt erzielten: Die Leistungen liegen in diesem Jahr ungefähr gleich viel über dem Durchschnitt, wie die Einschätzungen. Daraus resultiert eine durchschnittlich fast realistische Einschätzung, respektive die tiefste Überschätzung aller drei Messungen (Abb.4-2-8).

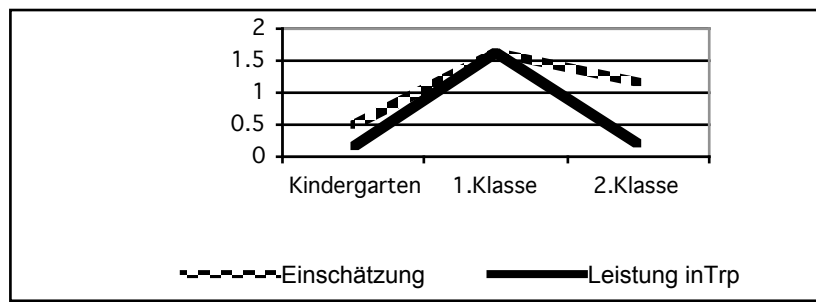


Abbildung 4-2-8: Abweichung der Einschätzung und Leistung vom erwarteten Mittelwert (=0) in Treppenstufen im Kindergarten, der ersten und zweiten Klasse

Zweitklässler überschätzen sich signifikant und verfügen über ein günstiges Selbstkonzept (Einschätzhöhe über dem Durchschnitt), weil sie sich bei den leistungsstärkeren Kindern ihres Jahrganges sehen. Die Erstklässler haben einerseits ein sehr hohes Selbstkonzept bezüglich ihrer mathematischen Fähigkeiten und gleichzeitig übertreiben sie in Bezug auf ihre Leistungsfähigkeit in der vorliegenden Untersuchung am wenigsten: Sie sind auch, gemessen an den Leistungsnormen des KAB, gute Rechnerinnen und Rechner. Im Kinder-

garten waren die Kinder demgegenüber vorsichtiger: Sie schätzten sich im sozialen Vergleich zwar ebenfalls signifikant über dem erwarteten Durchschnitt ein, allerdings am wenigsten ausgeprägt. Ihr Selbstkonzept bezüglich den mathematischen Fähigkeiten ist also vergleichsweise am tiefsten.

Zusammenfassend sei festgehalten, dass die untersuchten Kinder in allen Jahren über ein ausgeprägt positives Selbstkonzept (signifikant überdurchschnittliche Einschätzhöhe) bezüglich mathematischer Fähigkeiten verfügen. Gemessen an der sozialen Gruppe schätzen sich die Kinder im Kindergarten und in der zweiten Klasse signifikant höher über dem Durchschnitt ein, als gemessen an ihrer realen Leistung (Überschätzung). Sowohl Einschätzhöhe wie auch Überschätzung in Mathematik sind am Ende der zweiten Klasse höher als im Kindergarten. Die Befunde von Helmke (1991) zur Entwicklung der Einschätzhöhe, die als einzige diesen Fachbereich in einem Längsschnitt über diesen Zeitraum untersuchten, kamen zu einem anderen Ergebnis. Die Einschätzhöhe war in der zweiten Klasse niedriger als im Kindergarten, wenn auch nicht signifikant. Auch bei Helmke ergab sich ein diskontinuierlicher Verlauf mit der höchsten Einschätzhöhe in der ersten Klasse. Vergleicht man den Kindergarten hingegen mit der dritten, statt der zweiten Klasse, dann stieg der Wert bei Helmke ganz leicht an (vgl. Tab.1-2-1), von durchschnittlich 2.2 auf 2.3 (Skala: 5-bis+5).

4.2.7 Überprüfung der Zusammenhangshypothesen

Mit Zusammenhangshypothesen wird nun geprüft, ob sich die Genauigkeit der Leistungseinschätzung im Laufe der Schuljahre signifikant verbessert. Dann interessiert, ob sich Knaben oder Mädchen genauer einschätzen und ob sich entsprechende signifikante Unterschiede bei Stadt- und Landkindern ergeben. Die Korrelation von Einschätzung x Leistung überprüft, wie eng die Schätzungen der Kinder mit den Leistungen übereinstimmen (Tab.4-2-12).

Tabelle 4-2-12: Korrelationen von Einschätzung (auf der Treppenskala) mit der Leistung in Mathematik

	E 1	E 2	LKG	L1K	L2K
Einschätz g.Grp					
KG	.210**	.248**	.289**	0.234**	.198**
1.Kl.		.216**	.133	.181*	.167*
2.Kl.			.312**	.358**	.348**
Leistung					
KG.				.632**	.635**
1.Kl.					.650**

(Bedeutung der Abkürzungen, z.B.: E 1= Einschätzung auf der Treppe, in der 1.Klasse; L2K= Leistung in der 2.Klasse) * = $p < 0.05$; ** = $p < 0.01$)

Die Korrelationen von Einschätzung x Leistung sind in allen Jahren signifikant. In der ersten Klasse zeigt die Korrelation einen schwachen Effekt, im Kindergarten und der zweiten Klasse einen mittleren Effekt. Die Leistung erklärt im Kindergarten 7.9% und in der ersten Klasse 2.8% der Varianz. In der zweiten Klasse sind es 11.7%. Um die Korrelationen miteinander zu vergleichen, wurden die Koeffizienten in „Fishers Z“ umgerechnet. Anschliessend erfolgte für die klassenstufenweisen Vergleiche eine Signifikanzprüfung für abhängige Stichproben (Bortz, 1993, 205). Die Abnahme der Korrelation vom Kindergarten zu ersten Klasse ist nicht signifikant ($z=1.239$; $p>0.10$), der Unterschied der Korrelation des Kindergartens im Vergleich zur zweiten Klasse ist noch geringer, die leichte Zunahme der Korrelation also nicht signifikant. Die Zunahme der Korrelation von der ersten zur zweiten Klasse ist dagegen signifikant ($z=1.961$, $p < 0.05$). Die relativ niedrigen Korrelationen sind aber in allen drei Jahren etwas höher als bei Helmke (1991), in dessen Befunden sich erst in der zweiten Klasse ein signifikanter Effekt ergibt. Allerdings hat er vorher auch keine Leistungstests gemacht, sondern das Urteil der Mutter und der Lehrperson mit der Selbsteinschätzung korreliert.

Hypothese 3.1., wonach die Schätzgenauigkeit in Mathematik von Jahr zu Jahr zunimmt, kann nur teilweise bestätigt werden: Vom Kindergarten zur ersten Klasse nimmt sie (nicht signifikant) ab, vom ersten zum zweiten Schuljahr dann signifikant zu.

Hypothese 3.2., wonach die Schätzgenauigkeit in Mathematik vom Kindergarten zur zweiten Klasse zunimmt, kann nicht bestätigt werden.

Die Stabilität der Einschätzungen über die Jahre sagt etwas darüber aus, ob sich die Kinder von Jahr zu Jahr etwa gleich einschätzen. Die Stabilität der Einschätzungen nimmt tat-

sächlich kaum zu, die Korrelation zwischen den Einschätzungen KG.-1.Kl. beträgt $r=.210$ und 1.Kl.-2.Kl $r=.216$. Diese signifikanten Werte (je $p<0.01$) sind allerdings nicht sehr befriedigend. Sie sind etwas höher als bei Helmke, aber tiefer als bei Marsh et al. (1998).

In Bezug auf Geschlechtsunterschiede wissen wir, dass Knaben sich leistungsmäßig höher einschätzen als Mädchen, aber sie erbringen auch die signifikant besseren Leistungen. Das führt im insgesamt sogar dazu, dass sie sich im zweiten Schuljahr weniger überschätzen als die Mädchen. Heisst dies nun auch, dass ihre Schätzgenauigkeit grösser ist als diejenige der Mädchen?

Tabelle 4-2-13: Signifikanzprüfung der Korrelationen (Einschätzung x Leistung), getrennt nach Geschlecht

	Mädchen		Signifikanz (zweiseitig)	Knaben		Signifikanz (zweiseitig)
	Korrelation	F-Wert		Korrelation	F-Wert	
Kindergarten	.404	18.72	.000***	.184	3.504	.064
Erste Klasse	.135	1.78	.186	.182	3.442	.066
Zweite Klasse	.292	8.53	.004**	.363	15.133	.000***

Legende: * $p<.05$; ** $p<.01$; *** $p<.001$;

Mädchen: Signifikanz bei: $df_1=1$; $df_2=96$; Knaben: Signifikanz bei $df_1=1$; $df_2=100$

Nur in der zweiten Klasse erreicht die Korrelation bei den Knaben einen signifikanten Wert, der sogar noch etwas höher ist als bei den Mädchen. Am genauesten schätzen sich aber die Mädchen im Kindergarten ein, zu einem Zeitpunkt, indem sie noch gar keinen Mathematikunterricht haben.

Um zu untersuchen ob sich Knaben oder Mädchen genauer einschätzen, wurde das obige Verfahren mit den „Fisher's-Z“ für unabhängige Stichproben verwendet. Es ergaben sich folgende (z-standartisierte) Werte: Im insgesamt über die drei Jahre beträgt die Differenz Mädchen - Knaben $z=0.039$ und ist damit nicht signifikant. Allerdings ergaben sich aber Interaktionen: Im Kindergarten schätzen sich die Mädchen genauer ein (mit einem $z=1.699$, $p<0.05$) wäre der Unterschied sogar signifikant. In der ersten und der zweiten Klasse sind die Unterschiede klein (1.KL.: $z=0.047$; 2.Kl.: $z=0.081$).

Tabelle 4-2-14: Signifikanzprüfung der Korrelationen (Einschätzung x Leistung),
getrennt nach Stadt- und Landkindern

	Stadt		Signifikanz	Land		Signifikanz
	Korrelation	F-Wert	(zweiseitig)	Korrelation	F-wert	(zweiseitig)
Kindergarten	.369	14.459	.000***	.210	4.782	.031*
Erste Klasse	.202	3.898	.051	.175	3.298	.072
Zweite Klasse	.367	14.301	.000**	.345	14.075	.000***

Legende: * $p < .05$; ** $p < .01$; *** $p < .001$

Stadt: Signifikanz: $df\ 1 = 1$, $df\ 2 = 92$; Land: Signifikanz: $df\ 1 = 1$, $df\ 2 = 104$

Während sich im Kindergarten die städtischen Kinder genauer einschätzen als diejenigen auf dem Lande, ist die Schätzgenauigkeit in der zweiten Klasse bei Stadt- und Landkindern fast gleich. Die Unterschiede zwischen Stadt- und Landkindern sind insgesamt klein: Damit die Werte signifikant werden, müsste die Differenz der Korrelationskoeffizienten höher als 0.2 sein. Stadt- und Landkinder unterscheiden sich also bezüglich Schätzgenauigkeit nicht.

Bezüglich Geschlecht und Milieu ergeben sich keine signifikanten Unterschiede in der Schätzgenauigkeit. Am Ende des zweiten Schuljahres unterscheiden sich die Kinder in der Einschätzgenauigkeit, sowohl was das Milieu wie das Geschlecht betrifft, kaum noch.

4.2.8 Häufigkeitsverteilungen und Betrachtung von Extremschätzungen

Nun werden noch die Häufigkeitsverteilungen der Rohwerte für die beiden Variablen „Einschätzung“ und „Leistung“ dargestellt. Insbesondere interessieren dabei die Extremwerte, also beispielsweise Kinder, die sehr gute Leistungen erbracht haben. Wie hoch schätzen sich solche Kinder ein? Wie steht es um Kinder, die sehr schwache Leistungen erbringen? Diese Fragen sind für die Schule von hohem Interesse.

Insbesondere dort, wo die Maximal- und Minimalwerte der Skalen häufig genannt werden, kann es bei einer kleinen Skalenbreite zu „Ceiling“-Effekten und vor allem auch zu unerwünschten Regressionseffekten kommen. Die Korrelationen von Leistung und Einschätzung sind in unserer Stichprobe an zwei Orten belastet: Im Kindergarten und der 1.Klasse sind je in 5% der Fälle sehr hohe Einschätzungen (7 oder 8) gemacht worden, die dazu führen können, dass die zu diesen Einschätzungen gehörenden Leistungen durch Regressionseffekte belastet sind. Betrachtet man die gesamte Experimentalgruppe, dann wurden in den vorliegenden Befunden keine bedeutsamen ceiling –Effekte gefunden, entsprechend sind Regressi-

onseffekte klein. Da aber im Folgenden ausschliesslich Extremfälle betrachtet werden, können Regressionseffekte auftreten, weshalb für diese Gruppe nur sehr eingeschränkte Aussagen gemacht werden können. Zusätzlich ist die (durch die breite Skala ja gerade erhoffte), niedrige Anzahl an Fällen im Extrembereich so niedrig, dass keine quantitative Aussagen mehr möglich sind. Bei der Betrachtung der Extremfälle handelt es sich also vor allem um Einzelfallstudien, deren „Schicksal“ aus pädagogischer Sicht aber interessieren muss.

Tabelle 4-2-15: Häufigkeitsverteilung der Selbsteinschätzungen auf der Einschätztreppe. (schattiert die Einschätzungen in Bezug auf das eigene Ausbildungsjahr. Beispiel: -4 = „schwächstes Kind meines Ausbildungsjahres“)

Zahlenwert:	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
Kindergarten		2	1	8	10	3	3	43	53	9	29	6	20	4			9
1.Klasse					3	4	3	21	58	14	36	11	28	9	2	1	10
2.Klasse	1	1		1	3	1	6	22	63	12	43	13	20	8	2		4

Die Selbsteinschätzungen der Kinder in Tab.4-2-15 beschränken sich nicht nur auf den Rahmen der eigenen Klassenstufe (grau unterlegt). So schätzen sich beispielsweise im Kindergarten 11 Kinder leistungsmässig unterhalb ihres Ausbildungsjahres ein (-5 bis -8). Leicht kann man vermuten, dass eine so tiefe Einschätzung unrealistisch sei. Die beiden Kindergartenkinder mit der Einschätzung - 7 erbringen aber eine für ihr Alter sehr tiefe Leistung, sie erzielten 11 und 12 Rohwertpunkte. Nach den Normen des KAB-C entspricht dies einer unterdurchschnittlichen Leistung bei Kindern, die sogar ein Jahr jünger sind als Kindergartenkinder. Die tiefe Einschätzung ist damit ziemlich realistisch, denn der Wert -6 wäre der richtige. Dieses Beispiel zeigt, dass es sinnvoll ist, die Kinder ihre Leistung auf einer Skala einschätzen zu lassen, die mehr umfasst, als das eigene Ausbildungsjahr.

Tabelle: 4-2-16: Anzahl Personen, die einen bestimmten Rohwert (Punkte) im KAB erzielten (Fett: Mittelwerte KG; 1.Kl.;2.Kl.)

Rohwert	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Kindergarten	1	4		6	10	15	13	16	10	15	15	9	19	11	17	14
1.Klasse											1		2	9	20	25
2.Klasse															2	10

Rohwert	24	25	26	27	28	29	30	31	32	33	34	35	36
Kindergarten	11	7	5	2									
1.Klasse	45	44	23	9	7	6	4	2	2	1			
2.Klasse	17	35	30	26	23	20	12	6	5	7	5	1	1

Im Kindergarten erbringen schon 12.5% aller Kinder mindestens die Leistungen eines durchschnittlichen Schülers der ersten Klasse (24 Punkte, gemäss K-ABC-Norm), zwei von diesen Kindern sogar die Leistung eines durchschnittlichen Zweitklässlers. In der ersten Klasse beträgt der Anteil 11%, in der zweiten Klasse sind es 9.5 % der Kinder, die durchschnittliche Leistungen erbringen, wie sie von Kinder erwartet werden, die ein Schuljahr länger dabei sind. Der Anteil Kinder, die weit überdurchschnittliche Leistungen erbringt nimmt also ab.

Die folgende Tabelle (4-2-17) stellt dar, welche Leistungen in jedem Jahr von den Kinder erwartet werden können und wie diese zu beurteilen sind: Erbringt ein Kind im Kindergarten 22 Rohwertpunkte (nach KAB-C), dann ist das für Kindergartenkinder eine Spitzenleistung. Liegt der Wert bei 23 Punkten, dann ist das bereits eine durchschnittliche Leistung für Kinder, die ein Jahr älter sind. Punktzahlen, die rechts der grau unterlegten Felder sind, bedeuten, dass man leistungsmässig zum Durchschnitt oder sogar zur Spitze (+8) des nächsten Ausbildungsjahres gehört (entsprechendes gilt für Punktzahlen links der grauen Felder).

Tabelle 4-2-17: Verteilung der Leistungen über die Treppenstufen (-8 bis +8)
(links und rechts der Schattierung: Werte ausserhalb des eigenen Jahrgangs)

Kindergarten	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
Rohwertpunkte nach KAB	<11	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	>25
Häufigkeit der Leistung in %	2.5	3	5	7.5	6.5	8	5	7.5	7.5	4.5	9.5	5.5	8.5	7	5.5	3.5	3.5

Erste Klasse	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
Rohwertpunkte nach KAB	<12	13	16, 15	18, 17	19	20	21	22	23	24	25	26	27	28	29, 30	31	>31
Häufigkeit der Leistung in %				0.5		1	4.5	10	12.5	22.5	22	11.5	4.5	3.5	5	1	1.5

Zweite Klasse	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
Rohwertpunkte Nach KAB	<20	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	>34
Häufigkeit der Leistung in %				1	5	8.5	17.5	15	13	11.5	10	6	3	2.5	3.5	2.5	1

Der Prozentsatz von Kindern mit Leistungen, die unterhalb ihrer Altersgruppe liegen geht erfreulicherweise von Jahr zu Jahr zurück. Die Anzahl Kinder, die einen deutlichen Leistungsvorsprung haben, geht vom Kindergarten zu den Schuljahren ebenfalls zurück. Es werden nun zuerst diejenigen Kinder betrachtet, die sich sehr leistungsstark eingeschätzt haben (Tab.4-2-18).

Tabelle 4-2-18: Alle Kinder, die sich als „sehr gut, so gut wie der beste des nächsten Jahrganges“ eingeschätzt haben im Vergleich zur Gesamtstichprobe (Mittelwerte in Treppenstufen)

Anzahl	Einsch.KG	E 1.Kl.	E 2.Kl.	Leistg. KG	L. 1.Kl.	L. 2.Kl.
Kindergarten						
9 (4.5%)	8	3.11	3.55	4	6	2
1.Klasse						
10 (5%)	1.3	8	0.1	-1	1	2
2.Klasse						
4 (2%)	3.75	4.25	8	3	5	3
Alle (200)	.50	1.62	1.16	0	1	0

Die extrem hoch einschätzenden Kindergartenkinder erbringen innerhalb ihrer Klassenstufe sehr gute Leistungen, ein Jahr später durchschnittlich sogar Leistungen, die besser sind als bei den meisten Zweitklässlern (Treppenwert 6 in 1.Kl.). In der ersten Klasse sind es, mit einer Ausnahme, neun andere Kinder, die sich so hoch einschätzten. Deren tatsächliche Leistungen sind allerdings nur leicht über dem Durchschnitt ihrer Klassenstufe. Bei den vier extrem einschätzenden Kindern des zweiten Schuljahres waren die Leistungen im Jahr zuvor durchschnittlich sehr hoch, während sie nun klassenrelativ noch „gut“ sind. Überblickt man die Gruppe aller Kinder, die sich mindestens einmal zuoberst eingeschätzt haben, lässt sich festhalten, dass die Anzahl solcher Kinder in der zweiten Klasse nur noch sehr klein ist.

Tabelle 4-2-18: Kinder, die sich als „überhaupt nicht gut, bei den Schwächsten“ einschätzen

Kindergarten	Einsch.KG	E 1.Kl.	E 2.Kl.	Leistg. KG	L. 1.Kl.	L. 2.Kl.
1	-7	+1	-4	-7	-2	-3
2	-7	+1	+5	-6	-1	-2
3	-6	0	+4	+2	+1	0

2. Klasse	Einsch.KG	E 1.Kl.	E 2.Kl.	Leistg. KG	L. 1.Kl.	L. 2.Kl.
1	+2	0	-7	-5	-1	-3
2	-4	+8	-8	-5	0	-2
3	0	-3	-5	-5	2	0

Nun noch die Gruppe der extremen Unterschätzerinnen und Unterschätzer: Diese Gruppe ist deutlich kleiner, ja in der 1.Klasse schätzt sich kein Kind unterhalb seiner Vergleichsgruppe ein (-5 oder tiefer). Die tiefe Einschätzung im Kindergarten geht bei den ersten beiden Kindern mit einer tatsächlich sehr schwachen Leistung einher, demgegenüber unterschätzt sich das dritte Kind. Die schwachen Leistungen bei diesen Kindern setzten sich auch in den späteren Schuljahren fort. Die tiefe Selbsteinschätzung nahm aber erfreulicherweise ab. In zwei Fällen waren die Kinder am Ende des zweiten Schuljahres dann überoptimistisch, während sich ein Kind als sehr schwach im klassenrelativen Leistungsvergleich einstuft.

Bei den Zweitklässlern zeigt sich bei Kind „1“ eine stark sinkende Einschätzttendenz über die drei Messzeitpunkte: Die Leistungen dieses Kindes waren bisher noch in jedem Jahr unterdurchschnittlich, am Ende des zweiten Schuljahres ist dieses Kind leistungsmässig schwach. Bei Kind „2“ fällt die Unstetigkeit der Selbsteinschätzung auf, die mit konstant unterdurchschnittlichen Leistungen einher geht. Kind „3“ schätzt sich von Jahr zu Jahr deutlich schlechter ein und glaubt, am Ende des zweiten Schuljahres, knapp unterhalb der Klassenleistungsstreuung zu liegen. Seine Leistungen rechtfertigen eine so tiefe Beurteilung nicht, das Kind erbringt immerhin durchschnittliche Leistungen.

Immerhin ist die Gesamtzahl der sich sehr tief einschätzenden Kinder in unserer Stichprobe sehr niedrig, dies ist sicher ein Ergebnis der pädagogischen Bemühungen der Lehrkräfte. Allerdings ist sie in der zweiten Klasse wieder angestiegen. Auch wenn man den Regressionseffekt berücksichtigt, müssten diese Kinder tiefere Leistungen erbringen, um eine so tiefe Selbsteinschätzung zu rechtfertigen. Diese drei Kinder haben wenig Selbstvertrauen, sie glauben ja eigentlich, dass sie leistungsmässig in die erste Klasse gehören.

4.3 Ergebnisse im Sport: Monopedales Überhüpfen

Sport wurde über die beiden Variablen „rückwärts Balancieren“ und „monopedales Überhüpfen“ (im folgenden als „Hüpfen“ bezeichnet) operationalisiert und getrennt untersucht. Im Kapitel 4.3 werden die Befunde zum Hüpfen, im Kapitel 4.4 diejenigen zum rückwärts Balancieren präsentiert.

Beide Kapitel folgen demselben Aufbau: Zuerst werden die Befunde zur Leistung vorgestellt und auf allfällige Geschlechts- und Milieuunterscheide hin betrachtet. Dann erfolgt die Darstellung der Mittelwerte der Einschätzhöhe (gemessen an der Treppenskala), um mit Signifikanzprüfungen festzustellen, ob sich die Kinder signifikant über dem prognostizierten Durchschnitt einschätzen. Nachfolgend gilt es festzustellen, ob die Nullhypothesen, dass sich die Einschätzhöhe mit zunehmendem Alter der Kinder nicht signifikant verändert, verworfen oder beibehalten werden können. Die Befunde zur Einschätzhöhe werden jeweils mit einer ANOVA in einem einzigen Schritt untersucht. Dann folgt eine ANCOVA, mit der Leistungsunterschiede konstant gehalten werden sollen.

Für die Berechnung der Überschätzung konnte ein direkteres Verfahren als bei der Mathematik gewählt werden: Die Kinder schätzten ihre Leistungsfähigkeit neben der Treppenstufenskala auch an der sichtbaren Aufgabe ein. Jeweils vor der Leistungserbringung schätzten sie die Sprunghöhe oder die Strecke fehlerfreien Balancierens. Von dieser Einschätzung konnte die tatsächliche Leistung in Abzug gebracht, und die Differenz als Überschätzung, respektive Unterschätzung, bezeichnet werden. Damit konnten, im Unterschied zur Mathematik, die gleichen Skalen für die Leistung und die Einschätzung eingesetzt werden.

Ein weiteres Kapitel untersucht die Korrelationen zwischen Leistung und Einschätzung der Leistung. Zwei Fragen stehen dabei im Zentrum: Erreicht die Schätzgenauigkeit einen signifikanten Wert? Wird die Schätzgenauigkeit über die Jahre besser? Immer geht es auch hier um die Frage, ob es Unterschiede zwischen Stadt- und Landkindern, respektive Mädchen und Knaben, gibt.

4.3.1 Die Leistungen in monopedaalem Überhüpfen

Mittels einer dreifaktoriellen ANOVA mit den unabhängigen Faktoren Geschlecht und Milieu („between-subjects“) und der Klassenstufe („within-subjects“) wurden die Leistungen in einem einzigen Schritt berechnet.(Tab.4-3-1).

Tabelle: 4-3-1: ANOVA: Leistungen in Hüpfen

Quelle	Df	F	P
Klassenstufe	2	365.73	0.000**
Geschlecht	1	2.142	0.145
Milieu	1	0.070	0.792
Geschlecht x Milieu	1	2.379	0.125
Stufe x Geschlecht	2	3.551	0.031*
Stufe x Milieu	2	2.079	0.128
Stufe x Geschlecht x Milieu	2	0.317	0.721

Es ergibt sich eine signifikante Interaktion von Geschlecht und Stufe. Obwohl im Insgesamt kein signifikanter Geschlechtsunterschied besteht, gibt es offenbar auf jeder Klassenstufe geschlechtsspezifische Leistungsdifferenzen, was Abb.4-3-1 bestätigt:

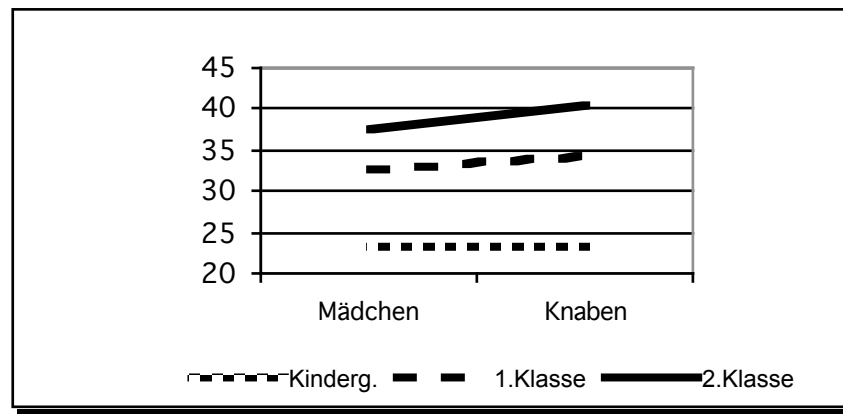


Abbildung 4-3-1: Leistungen von Knaben und Mädchen im Hüpfen (in cm).

Die Leistungsunterschiede scheinen klein. Die klassenstufenweise Prüfung ergibt im Kindergarten und der ersten Klasse nicht signifikante Unterschiede, erst im zweiten Schuljahr ergibt sich ein signifikanter Geschlechtsunterschied (2.85 cm, $p < 0.05$; Tab.4-3-2).

Tabelle 4-3-2: Durchschnittliche Leistungen im Hüpfen (in cm) von Knaben und Mädchen, sowie UNIANOVAs für jede Klassenstufe zur Prüfung von Geschlechtsunterschieden in der Leistung

	Knaben	Mädchen	Df	F	P
Kindergarten	23.20 cm	23.10 cm	1	0.003	0.956
Erste Klasse	34.20 cm	32.55 cm	1	1.655	0.200
Zweite Klasse	40.50 cm	37.65 cm	1	6.156	0.014*

(* sig.: $p < 0.05$)

Die Betrachtung der Leistungen zeigt auf allen Stufen höhere Werte für die Knaben, die Unterschiede nehmen über die drei Stufen zu, die Leistungsentwicklung der Geschlechter ist signifikant nicht parallel.

Dieses Ergebnis wurde nicht erwartet, denn aus den Normtabellen des Körperkoordinationstests „KTK“ (Kiphard & Schilling, 1974) sind die Mädchen im Kindergartenalter signifikant ($p < 0.01$) und im ersten Schuljahr tendenziell ($p = 0.063$) besser, während dann in der zweiten Klasse die Knaben etwas besser (aber nicht signifikant: $p = 0.245$) sind. Allerdings sind die Kinder der KTK-Normtabellen durchschnittlich vier Monate jünger, die Leistungsüberlegenheit der Knaben in den vorliegenden Befunden (leichte Tendenz bereits in der ersten Klasse) erklärt das teilweise. Dass Knaben bereits im Kindergarten leicht besser sind, fand auch Valkanover (2003).

Für die vorliegende Untersuchung wurde speziell darauf geachtet, Aufgaben auszuwählen, die sowohl für Knaben und Mädchen etwa gleich schwierig sind. In der KTK-Normierung war dies auch der Fall, den dort sind die Knaben nur im vierten Schuljahr signifikant besser ($p = 0.01$), in allen anderen Schuljahren sind die Unterschiede nicht signifikant (Daten bis ins achte Schuljahr). Mit unseren Daten stimmt das nicht überein, allerdings ergeben sich ja signifikante Leistungsvorteile für die Knaben erst im zweiten Schuljahr. Auf die Einschätzungen der Leistungen konnte das im zweiten Schuljahr (auf der Treppe und an der Aufgabe) kaum Einfluss haben, denn diese Einschätzungen erfolgten ja vor der Leistungserbringung.

In der Tab.4-3-1 wurde eine Interaktion Stufe x Geschlecht festgestellt. Die nun erfolgte Analyse zeigte, dass es sich um eine ordinale Interaktion handelt, sodass der Haupteffekt trotzdem interpretiert werden kann. Die Leistung nimmt insgesamt signifikant zu, die Kontrastanalyse ergibt für jeden Vergleich signifikante Differenzen (mit jeweils $p < 0.001$), die Leistung nimmt vom Kindergarten zur ersten Klasse um 10.3 cm zu, von der ersten zur zweiten Klasse um 5.6 cm. Im Kindergarten überspringen die Kinder eine durchschnittliche Höhe von 23 cm, in der zweiten Klasse bereits 38.9 cm.

Zusammenfassend kann festgehalten werden, dass die Leistungen der Kinder im Hüpfen von Jahr zu Jahr signifikant zunehmen, wobei die Leistungen der Knaben höher sind. Sie erbringen aber nur im zweiten Schuljahr signifikant bessere Leistungen als Mädchen. Sehr gering sind die Leistungsunterschiede zwischen Stadt- und Landkindern (vgl. Tab.4-3-1), diesbezüglich fanden sich auch in den KTK-Normen keine signifikanten Leistungsunterschiede.

4.3.2 Höhe der Einschätzung auf der Treppenskala

Die Gesamtzahl der Kinder, die alle drei Jahre die Aufgabe bewältigt haben, liegt etwas unterhalb derjenigen in Mathematik, weil einige Kinder über längere Zeit verletzt waren und nicht nachuntersucht werden konnten (Tab.4-3-3).

Tabelle 4-3-3: Lageparameter und Streuungsmasse der Selbsteinschätzung auf der Treppenskala im Hüpfen vom Kindergarten zur zweiten Klasse

	Schätzung der Leistung, Kinderg.	Schätzung der Leistung, 1.Klasse	Schätzung der Leistung, 2.Klasse
N	194	194	194
Arithm. Mittel	.57	0.73	0.80
Median	0	0	0
Modus	0	0	0
Std. Abweichung	2.953	2.420	1.983
Varianz	8.723	5.858	3.930
Minimum	-8	-6	-5
Maximum	8	8	8

(Mittelwert .57 bedeutet eine Abweichung der Einschätzung um etwas mehr als eine halbe Treppenstufe nach oben)

Im Durchschnitt liegt das arithmetische Mittel in allen Jahren über dem Wert 0. Im Kindergarten beträgt die Abweichung der Einschätzung im t-Test für eine Stichprobe $t(193) = 2.698$, $p < 0.01$ und ist damit signifikant. In der ersten Klasse beträgt sie $t(193) = 4.183$, $p < 0.01$, in der 2. Klasse $t(193) = 5.649$, $p < 0.01$ und ist damit ebenfalls signifikant, die Mittelwerte sind also alle positiv signifikant abweichend.

Die Hypothese 1.1. wird damit bestätigt, die Kinder schätzen sich in „Hüpfen“ durchschnittlich im Kindergarten, der ersten und der zweiten Klasse signifikant über dem erwarteten Mittelwert (=0) ein.

Median und Modus liegen durchwegs bei 0, die grösste Anzahl Einzeleinschätzungen der Leistung findet sich also genau in der Skalenmitte. Sie steigt zuerst von 62 auf 94 und sinkt dann auf 70 (Abb.4-3-2).

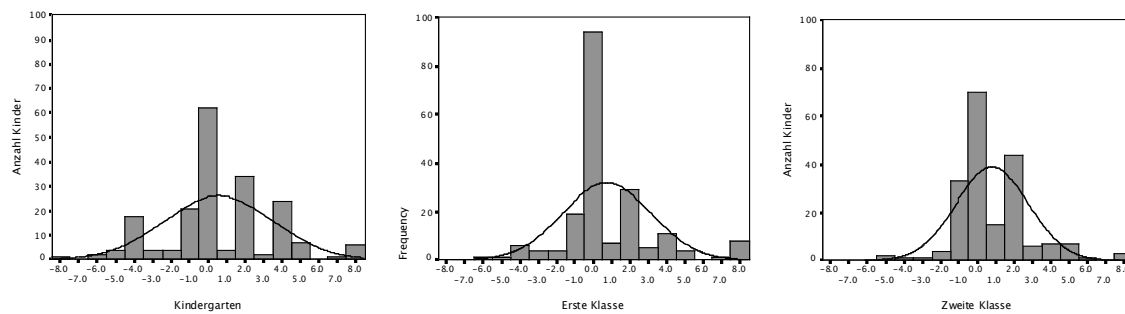


Abbildung 4-3-2: Einschätzung der eigenen Leistungsfähigkeit auf der Treppenskala in Hüpfen

Über dem Median schätzen sich im Kindergarten 78 (40.2%) der Kinder ein, unterhalb sind es 54 (27.8%). In der ersten Klasse liegen 65 (33.5%) der Kinder über dem Median und 35 (18.0%) darunter, während es in der zweiten Klasse 83 (42.8%) sind, die über, respektive 41 (21.1%) die unter dem Median liegen. Damit verändert sich das Verhältnis von über-, respektive unter dem Median einschätzenden Kindern über die drei Messzeitpunkte: Der Anteil der sich über dem Median Einschätzenden nimmt von 59.1% im Kindergarten, auf 65% in der ersten, respektive 66.7 % in der zweiten Klasse zu, dennoch ist die Regression nicht signifikant.

Der Vergleich der Standardabweichungen, respektive der Varianzen ergibt durchwegs signifikante Unterschiede: KG-1.Kl.: ($F(1.35)=1.489$, $p<0.05$), KG-2.Kl.: ($F(1.53)=2.219$, $p<0.01$) und 1.Kl.-2.Kl.: ($F(1.53)=1.1775$, $p>0.01$). Gemäss Bortz (1993, 263) sind diese Varianzdifferenzen vernachlässigbar, da die Stichprobengrößen identisch und hoch sind, die Verteilungen zudem symmetrisch. Die Unterschiede interessieren substanziell, die Anzahl Einschätzungen an der Klassenspitze oder als Klassenschwächste ist in der ersten Klasse deutlich tiefer als im Kindergarten. Auch die Anzahl der Einschätzungen der Art, dass man eigentlich leistungsmässig zu den um ein Jahr älteren oder jüngeren Kindern gehört, wird kleiner. Die Kinder sehen sich ein Jahr später, in der zweiten Klasse, leistungsmässig zwischen leicht unter dem Klassendurchschnitt („manchmal nicht so gut“) und bei den besten Zweitklässlern („sehr gut“). Das stützt die Vermutung, dass sich die Kinder im Vergleich zum Kindergarten in den Schuljahren beim Einschätzen zunehmend an der eigenen Klasse orientieren.

4.3.3 Befunde zur Einschätzhöhe unter Berücksichtigung von Klassenstufe, Geschlecht und Milieu

Mittels einer dreifaktoriellen ANOVA mit den unabhängigen Faktoren Geschlecht und Milieu („between-subjects“) und der Klassenstufe („within-subjects“) wurden die Befunde in einem einzigen Schritt berechnet. (Tab.4-3-4).

Tabelle 4-3-4: ANOVA: Einschätzung der Leistung anhand der Treppenskala in Hüpfen

Quelle	Df	F	P
Klassenstufe	2	0.344	0.689
Geschlecht	1	7.811	0.006**
Milieu	1	1.015	0.315
Geschlecht x Milieu	1	0.072	0.789
Stufe x Geschlecht	2	0.104	0.885
Stufe x Milieu	2	1.180	0.306
Stufe x Geschlecht x Milieu	2	1.492	0.223

Die Interaktionen sind nicht signifikant, die Haupteffekte sind direkt interpretierbar. Nur der Geschlechtseffekt ist signifikant, Knaben schätzen sich signifikant höher ein als Mädchen (1.03 versus 0.377 Treppenstufen über dem Durchschnitt). Auch wenn die durchschnittliche Einschätzhöhe (wie oben diskutiert) über die Messzeitpunkte hinweg zunimmt, so liegen die Unterschiede doch innerhalb der Zufallsschwankung. Die Unterschiede zwischen Stadt- und Landkindern sind ebenfalls nicht signifikant. Die Klassenstufen unterscheiden sich nicht signifikant.

Die Hypothese 2.3., wonach die Einschätzhöhe im Monopedalen Überhüpfen,

a) im Kindergarten im Vergleich zum ersten Schuljahr gleich hoch ist, wird nicht verworfen,

b) in der ersten Klasse im Vergleich zur zweiten Klasse gleich hoch ist, wird nicht verworfen.

die in beiden Fällen gemessene Zunahme der Einschätzhöhe ist nicht signifikant.

Die Hypothese 2.4., wonach die Einschätzhöhe im Monopedalen Überhüpfen im Kindergarten gleich hoch ist wie in der zweiten Klasse, kann nicht verworfen werden. Die Einschätzhöhe im Kindergarten unterscheidet sich nicht signifikant von derjenigen in der zweiten Klasse.

In Tab.4-3-4 ergab sich ein signifikanter Geschlechtseffekt. Der Vergleich zeigt, dass sich die Knaben immer höher einschätzen als die Mädchen (Abb.4-3-3).

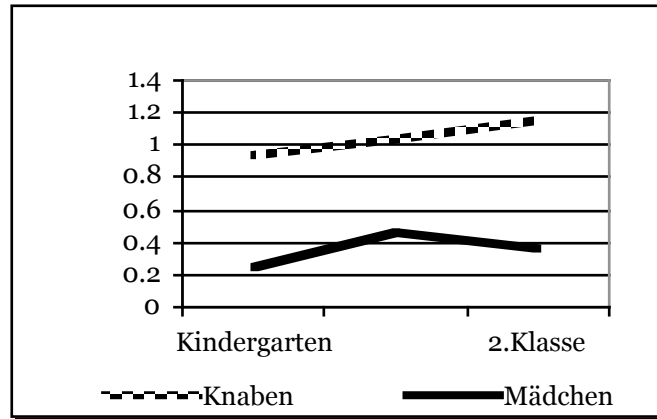


Abbildung 4-3-3: Durchschnittliche Abweichung der Einschätzung von Knaben und Mädchen (1= eine Treppenstufe über der Treppenmitte)

Die Hypothese 4.2., wonach sich die Knaben im Sport signifikant höher einschätzen als die Mädchen, kann für das „Monopedale Überhüpfen“ bestätigt werden.

In Tabelle 4-3-4 ergab sich kein signifikanter Milieueffekt

Die Hypothese 5.2., wonach sich die urbanen Klassen im Sport höher einschätzen als die ruralen kann für das Monopedale Überhüpfen nicht bestätigt werden.

Mit klassenstufenweisen UNIANCOVAs werden nun jahrgangsweise die Einschätzungen auf Geschlechts- und Milieuunterschiede geprüft und zwar unter Konstanthaltung der Leistung. Bei dieser Prüfung zeigt sich, dass die Leistung in der ersten Klasse einen signifikanten Einfluss auf die Einschätzung hat ($F=6.409$, $p<0.05$). Im Kindergarten ist der Einfluss mit $F=0.069$, $p>0.10$ und im zweiten Schuljahr mit $F=0.979$, $p>0.10$ dagegen nicht signifikant (Tab.4-3-5).

Tabelle 4-3-5: Klassenstufenweise UNIANCOVA im Hüpfen, mit Geschlecht und Milieu als unabhängigen Faktoren und der Leistung als Kovariate

Quelle	Df	F	P
Kindergarten			
Geschlecht	1	2.292	0.132
Milieu	1	0.057	0.811
Geschlecht x Milieu	1	0.705	0.403
Erste Klasse			
Geschlecht	1	2.238	0.134
Milieu	1	3.125	0.071
Geschlecht x Milieu	1	0.704	0.403
Zweite Klasse			
Geschlecht	1	6.971	0.009**
Milieu	1	0.000	0.989
Geschlecht x Milieu	1	1.327	0.251

Die Interaktionen sind nicht signifikant. Die Haupteffekte sind darum direkt diskutierbar. Das Milieu hat in keinem Jahr einen signifikanten Effekt. Hingegen hat das Geschlecht in der zweiten Klasse einen signifikanten Effekt (0.765 Treppenstufen sind die Knaben besser, sig. mit $p < 0.01$). Ein Vergleich der einzelnen Ausbildungsjahre in Abb.4-3-4 zeigt einen nicht signifikanten Anstieg der Einschätzungshöhe bei den Knaben. Das Ansteigen der Einschätzung bei den Mädchen und der anschließende Abfall sind nicht signifikant ($p > 0.10$).

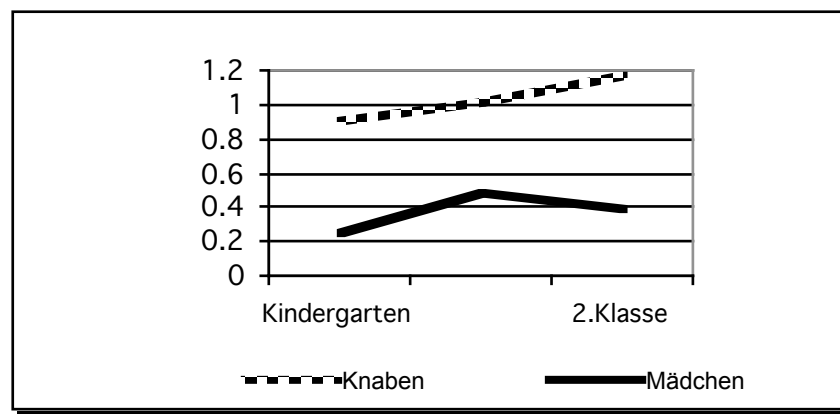


Abbildung 4-3-4: Höhe der Einschätzung im Hüpfen in Treppenstufen über dem Durchschnitt, im Kindergarten, der ersten und zweiten Klasse, unter Kovarianz der Leistung

Im Hüpfen konnten die Kinder neben der Leistungseinschätzung auf der Treppe zusätzlich einschätzen, wie viele Schaumgummis sie überhüpfen werden. Während sie im Balancieren rückwärts wussten, dass im Durchschnitt ihres Altersjahrganges ein Kind bis in die Mitte der Balancierstrecke kommt (gleich bleibende Schwierigkeit der Aufgabe), nahm die zu erwartende Durchschnittsleistung beim Hüpfen von Jahr zu Jahr zu. In jedem Jahrgang erhielt das Kind deshalb entsprechende Hinweise, wieviele Schaumgummis seine Klassenstufe überspringt (vgl. die genauen Instruktionen im Anhang). Durchschnittlich überspringt ein Kind gemäss Instruktion im Kindergarten 4 Schaumgummis, im ersten 5 und im zweiten Schuljahr 6 Schaumgummis. Im t-Test erwies sich die durchschnittliche Abweichung der Einschätzung von der tatsächlichen Leistung über die drei Klassenstufen als signifikant ($p < 0.01$). Die Kinder wichen in ihren Einschätzungen im Kindergarten mit 1.15, in der ersten Klasse mit 1.08 und in der zweiten Klasse mit 1.06 Schaumgummis ab und zwar so, dass sich auf jeder Klassenstufe eine signifikante, über dem Durchschnitt liegende Einschätzhöhe ergab. Wie die nahe beieinander liegenden Werte schon andeuten, zeigt die Kontrastanalyse dabei keine signifikanten Unterschiede zwischen den jeweiligen Werten.

In Mathematik konnte die Einschätzung nur über die soziale Bezugsnorm, also über die Treppenstufenskala, gemessen werden. Im Hüpfen konnte die Kinder vor der Leistungserbringung noch zusätzlich eine Einschätzung an der Aufgabe vornehmen und die vermutete Maximalhöhe schätzen, die sie sich zutrauten. Die Befunde in Tab.4-3-6 können direkt mit denjenigen aus Tab.4-3-4 verglichen werden. In beiden Tabellen finden sich signifikante Befunde an der gleichen Stelle, nämlich beim Geschlecht.

Tabelle 4-3-6: Dreifaktorielle ANOVA zur Einschätzung der Leistung an der Aufgabe „Hüpfen“ (abhängige Variable), sowie Stufe, Geschlecht und Milieu (unabhängige Variablen).

Quelle	Df	F	P
Klassenstufe	2	0.199	0.812
Geschlecht	1	11.218	0.001
Milieu	1	1.178	0.279
Geschlecht x Milieu	1	0.227	0.137
Stufe x Geschlecht	2	0.483	0.610
Stufe x Milieu	2	0.419	0.651
Stufe x Geschlecht x Milieu	2	0.568	0.501

Interaktionen sind keine vorhanden. Von den Haupteffekten ist der Geschlechtseffekt signifikant. Die Knaben schätzen sich über die drei Jahre gesehen um durchschnittlich 6.70 cm, die Mädchen um 4.27 cm über dem Durchschnitt ein. Der Klassenstufeneffekt ist nicht

signifikant, die Einschätzungen unterscheiden sich nicht voneinander. Die Einschätzhöhe bleibt über die Klassenstufen hinweg konstant.

Die Hypothese 2.3. wonach die Einschätzhöhe (gemessen an der Aufgabe) zwischen allen Jahren gleich hoch bleibt, wird nicht verworfen.

Hypothese 2.4., wonach die Einschätzhöhe (gemessen an der Aufgabe) im Kindergarten gleich hoch ist wie in der zweiten Klasse, wird nicht verworfen.

Fazit: Sowohl die an der sozialen, wie die an der Aufgabe direkt gemessenen Einschätzungen weichen von Jahr zu Jahr nur innerhalb der Zufallschwankungen voneinander ab, die Einschätzhöhe bleibt also vom Kindergarten bis zur zweiten Klasse gleich hoch. Die Knaben schätzen sich in beiden Fällen signifikant höher ein als die Mädchen. Es ergeben sich keine signifikanten Milieuunterschiede.

4.3.4 Über- oder Unterschätzung im monopedalen Überhüpfen

Zieht man nun von der Einschätzung (an der Aufgabe) die tatsächlich übersprungene Anzahl Schaumgummis ab, dann kann der Rest als Über-, respektive Unterschätzung bezeichnet werden. Ein erster Blick auf Tab.4-3-7 zeigt, dass sich die Kinder in den beiden Schuljahren im Durchschnitt unterschätzten.

Tabelle 4-3-7: Lageparameter und Streuungsmasse der Überschätzung im Hüpfen vom Kindergarten zur zweiten Klasse

	Überschätzung der Leistung, Kinderg.	Überschätzung der Leistung, 1.Klasse	Überschätzung der Leistung, 2.Klasse
N	194	194	194
Arithm. Mittel	0.55	-0.56	-0.74
Median	0	0	-1
Modus	0	0	0
Std. Abweichung	2.244	1.968	1.941
Varianz	5.036	3.874	3.766
Minimum	-4	-6	-6
Maximum	8	5	4

(Mittelwert .55 bedeutet eine Überschätzung um etwas mehr als eine halbe Treppenstufe nach oben)

Im Durchschnitt haben sich die Kinder im Kindergarten signifikant überschätzt, der t-Test ergibt $t(193)=3.423$, $p<0.01$. In der ersten und zweiten Klasse ergibt sich dagegen je eine signifikante Unterschätzung (1.Kl.: $t(193)=-3.976$, $p<0.01$; 2.Kl.: $t(193)=-5.304$, $p<0.01$).

Die Varianzen der verschiedenen Klassenstufen unterscheiden sich nicht signifikant. Median und Modus fallen, mit Ausnahme der zweiten Klasse, auf den gleichen Wert.

Tabelle 4-3-8: Dreifaktorielle ANOVA zur Überschätzung der Leistung in Hüpfen als abhängiger Variable, sowie Stufe, Geschlecht und Milieu als unabhängigen Variablen

Quelle	Df	F	P
Klassenstufe	2	28.456	0.000
Geschlecht	1	0.649	0.421
Milieu	1	0.990	0.321
Geschlecht x Milieu	1	0.277	0.599
Stufe x Geschlecht	2	1.328	0.266
Stufe x Milieu	2	1.758	0.174
Stufe x Geschlecht x Milieu	2	0.568	0.567

In Tab. 4-3-8 ergeben sich keine Interaktionen. Von den Haupteffekten ist der Klassenstufeneffekt signifikant. Die Kontrastanalyse des Haupteffekts zeigt, dass der Mittelwert vom Kindergarten sowohl zur ersten, wie zur zweiten Klasse, signifikant ($p < 0.01$) abnimmt. Der Unterschied zwischen erster und zweiter Klasse ist nicht signifikant (Differenz: 0.146 Treppenstufen, $p > 0.10$).

Knaben und Mädchen unterscheiden sich nicht signifikant in der Überschätzung. Dieses Ergebnis kontrastiert mit demjenigen bei der Einschätzung auf der Treppenskala und an der Aufgabe: Dort schätzten sich die Knaben signifikant höher ein. Diese höheren Einschätzungen sind also insofern gerechtfertigt, als die Knaben auch höhere Leistungen (in der zweiten Klasse sogar signifikant) erbringen als die Mädchen (vgl. Tab.4-3-2). Stadt- und Landkinder unterscheiden sich in der Überschätzungstendenz nicht signifikant.

Die Ergebnisse zur Überschätzung und deren Entwicklung über die drei Jahre sind allerdings nur bedingt aussagekräftig, weil die Instruktion nicht ganz stimmte: in jedem Jahr erzielten die untersuchten Kinder eine durchschnittliche Leistung, die um eine Schaumgummimatte höher lag, als aufgrund der KTK-Eichung zu erwarten gewesen wäre. Als die Kinder die Aufgabe einschätzten wurde Ihnen beispielsweise im ersten Schuljahr gesagt, dass die durchschnittlichen Erstklässler über 5 Matten springen (gemäss KTK-Norm), tatsächlich sprangen die Kinder der vorliegenden Untersuchung aber durchschnittlich über mehr als 6 Matten. Kinder, die sich durchschnittlich einschätzen wollten, schätzten also bei 5 Matten ein: Falls sie tatsächlich eine durchschnittliche Leistung erbrachten, sprangen sie dann über 6 Matten. Offensichtlich haben sich die diesbezüglichen Leistungen der Kinder in den vergan-

genen 20 Jahren seit der Eichung verbessert, sodass die Normierung im KTK von zu tiefen Leistungen ausgeht. Die gleiche Feststellung machte auch Valkanover (2003).

Die Instruktion hätte also in der ersten und zweiten Klasse von einem höheren Leistungsdurchschnitt ausgehen müssen, um „richtig“ zu sein. Würde man die Einschätzung an der Aufgabe in der ersten und zweiten Klasse korrigieren, das heißt den Schätzwert um durchschnittlich eine Matte erhöhen, dann wären die Mittelwerte der Ueber-, respektive Unterschätzung vom Kindergarten bis zur zweiten Klasse ganz leicht sinkend. Die Unterschiede in der Über-, respektive Unterschätzung von Jahr zu Jahr wären aber nicht mehr signifikant (Mittelwerte angepasst, in Treppenstufen: KG: 0.55; 1.Kl.: 0.44; 2.Kl.: 0.26). Die Überschätzung würde über die Jahre nicht signifikant sinken.

Die Nullhypothese 1.2. wird verworfen: Im Kindergarten überschätzen sich die Kinder im monopedalen Überhüpfen signifikant. In der ersten und zweiten Klasse unterschätzen sich die Kinder signifikant.

Hypothese 4.3., wonach sich Knaben mehr überschätzen als Mädchen, kann nicht bestätigt werden, der Unterschied ist nicht signifikant.

Hypothese 5.2. wird nicht bestätigt, urbane Kinder überschätzen sich nicht signifikant stärker als rurale.

Aufgrund der obigen Ausführungen stehen die Aussagen zur Hypothese 1.2. auf wackligen Füßen, da sie durch die Instruktionsprobleme vermutlich erheblich belastet sind.

4.3.5 Überprüfung der Zusammenhangshypothesen

Mit der Überprüfung von Zusammenhangshypothesen gilt es herauszufinden, ob sich die Genauigkeit der Leistungsschätzung im Laufe der Schuljahre signifikant verändert und ob sich Knaben oder Mädchen, Stadt- oder Landkinder signifikant unterschiedlich einschätzen. Die Korrelation von Einschätzung x Leistung überprüft, wie eng die Schätzungen der Kinder mit der tatsächlichen Leistungsfähigkeit übereinstimmen. Beim Hüpfen ist die Einschätzung an zwei Massstäben erfolgt, an der Aufgabe direkt und auf der Treppe.

Tabelle 4-3-9: Korrelationen von Einschätzung (auf der Treppe und an der Aufgabe) x Leistung im Hüpfen

	EG1	EG2	EAK	EA1	EA2	LKG	L1K	L2K
Einschätzg. Trp.								
KG	.254**	.124	.041	.202**	.118	-.031	0.21	-.012
1.Kl.		.346**	.170	.302**	.272*	.174*	.183*	.12
2.Kl.			.084	.218**	.320**	.184*	.232**	.103
Einschätzg. Aufg.								
KG				.149*	.082	.141*	.153*	.101
1.Kl.					.217**	.112	.238**	.181*
2.Kl.						.185*	.328**	.312**
Leistung								
KG.							.547**	.491**
1.Kl.								.647**

(Bedeutung der Abkürzungen, z.B.: EG1= Einschätzung auf der Treppe, in der 1.Klasse;
EAK= Einschätzung an der Aufgabe im Kindergarten; L2K= Leistung in der 2.Klasse)

Die Genauigkeit der Schätzung auf der Treppe (Einschätzung Trp x Leistung) erreicht nur in der ersten Klasse einen schwachen signifikanten Wert. Deutlich leichter gefallen ist den Kindern die Einschätzung an der Aufgabe selbst: Die Korrelationen sind in allen Fällen signifikant und steigen kontinuierlich an (von 0.141 über 0.238 auf 0.312), ab der ersten Klasse ergibt sich ein mittlerer Korrelationseffekt, im Kindergarten ein schwacher. Um die Korrelationen miteinander zu vergleichen, wurden die Koeffizienten in „Fisher's Z“ umgerechnet. Anschließend erfolgte für die klassenstufenweisen Vergleiche eine Signifikanzprüfung für abhängige Stichproben (nach Bortz, 1993, 205). Die Korrelationsdifferenz KG-1.Kl. beträgt 0.097. Diese Zunahme ist nicht signifikant ($z=1.005$; $p>.010$). Die Korrelationsdifferenz zwischen den beiden Schulklassen ist noch kleiner (0.076) und damit auch nicht signifikant. Zwischen Kindergarten und zweiter Klasse beträgt die Differenz 0.171, was einem signifikanten Unterschied entspricht ($z=2.109$; $p<0.05$). Die Genauigkeit nimmt vom Kindergarten zur zweiten Klasse linear zu, die Regression ist signifikant ($F=165.8$, $p<0.05$).

Hypothese 3.3., wonach die Schätzgenauigkeit in monopedalem Überhüpfen von Jahr zu Jahr zunimmt, kann nicht bestätigt werden: Vom Kindergarten zur ersten Klasse nimmt sie (nicht signifikant) zu, vom ersten zum zweiten Schuljahr nimmt sie auch zu (aber eben nicht signifikant).

Hypothese 3.4., wonach die Schätzgenauigkeit in monopedalem Überhüpfen vom Kindergarten zur zweiten Klasse signifikant zunimmt, kann bestätigt werden.

Die Einschätzungen auf der Treppe korrelieren mit der Leistung nur in der ersten Klasse schwach signifikant: Das Einschätzen der eigenen Leistungsfähigkeit ist damit in dieser Sportaufgabe wider Erwarten wenig genau ausgefallen. Dies im Gegensatz zur Mathematik, wo die Einschätzungen mit der Leistung in jedem Jahr signifikant korrelieren. Besonders im Kindergarten konnte ein anderes Ergebnis erwartet werden, weil dort ja die Vergleichsmöglichkeiten mit gleichaltrigen Mitschülern durch den gemeinsamen Sportunterricht Rückschlüsse auf die eigene Leistungsfähigkeit zu lassen, während es noch keinen gemeinsamen Mathematik- oder Rechenunterricht gibt.

Die Stabilität einer Einschätzung sagt etwas darüber aus, ob sich Kinder auf einer Skala von Jahr zu Jahr ähnlich oder sehr unterschiedlich einschätzen. Der zunehmende Schulunterricht müsste beispielsweise dazu führen, dass Kinder sich in Bezug zu ihren Leistungen und zu ihren Mitschülerinnen und -schülern von Jahr zu Jahr ähnlicher einschätzen, sofern die Bezugsgruppe gleich bleibt. Andererseits werden einzelne Kinder im Verhältnis zu den anderen leistungsstärker oder schwächer, weil sie beispielsweise ein höheres Potential haben oder eine bessere Aufgabenbetreuung erfahren, usw.

Die Stabilität der Einschätzungen sagt aber auch etwas über die Qualität des Messinstrumentes: Wird dieses von den Kindern nicht verstanden, dann werden sie willkürlich raten und aufgrund des Jahresabstandes sind dann grosse Schwankungen in den Einschätzungen möglich, also sehr tiefe Stabilitäten.

Die Stabilität der Einschätzungen im Hüpfen ist nicht sehr hoch, aber doch signifikant und über die beiden Jahresintervalle zunehmend: So korrelieren die Einschätzungen (auf der Treppenstufenskala) KG-1.Kl. mit $r=0.254$ und 1.Kl.-2.Kl. mit $r=0.346$. Die Stabilitäten bei der Einschätzung an der Aufgabe dagegen sind geringer, aber immer noch signifikant: KG-1.Kl. mit $r=0.149$ und 1.Kl.-2.Kl. mit $r=0.217$.

In Tabelle 4-3-10 werden die unterschiedlichen Schätzgenauigkeiten von Knaben und Mädchen auf Signifikanz geprüft.

Tabelle 4-3-10: Korrelationen (Einschätzung an der Aufgabe x Leistung), getrennt nach Knaben und Mädchen

	Mädchen		Signifikanz	Knaben		Signifikanz
	Korrelation	F-Wert	(zweiseitig)	Korrelation	F-Wert	(zweiseitig)
Kindergarten	.169	2.711	.103	.120	1.431	.234
Erste Klasse	.125	1.470	.229	.294	9.286	.003**
Zweite Klasse	.361	13.771	.000**	.253	6.637	.011*

Legende: *p<.05 ; **p<.01 ; Mädchen: Sign. bei: df 1= 1; df 2 = 96; Knaben: Sign bei df 1= 1; df 2 = 100

In der ersten und zweiten Klasse erreicht die Korrelation bei den Knaben einen signifikanten Wert. Am genauesten schätzen sich aber die Mädchen in der zweiten Klasse ein. Um zu untersuchen ob sich Knaben oder Mädchen genauer einschätzen, wurde das obige Verfahren mit den „Fisher's-Z“ für unabhängige Stichproben verwendet. Es ergaben sich folgende (z-standardisierte) Werte: Im insgesamt über die drei Jahre beträgt die Differenz Mädchen–Knaben $z= 0.003$ und ist damit nicht signifikant. Im Kindergarten und der zweiten Klasse schätzen sich die Mädchen genauer ein als die Knaben. Am grössten ist die Differenz zwischen den Geschlechtern aber in der ersten Klasse. Mit einem $z=1.213$ im t-Test ist der Unterschied aber nicht signifikant. In keinem Jahr unterscheiden sich Mädchen und Knaben signifikant bezüglich der Schätzgenauigkeit im Hüpfen.

Bei den Landkindern nimmt die Schätzgenauigkeit laufend zu und erreicht in der zweiten Klasse einen signifikanten Wert und eine mittlere Korrelation. Die Stadtkinder schätzen sich in der ersten und zweiten Klasse genauer ein als im Kindergarten (Tab.4-3-11).

Tabelle 4-3-11: Korrelationen (Einschätzung x Leistung), getrennt nach Milieu, von Stadt- und Landkindern

	Stadt		Signifikanz	Land		Signifikanz
	Korrelation	F-Wert	(zweiseitig)	Korrelation	F-Wert	(zweiseitig)
Kindergarten	.191	3.464	.066	.096	0.922	.339
Erste Klasse	.300	9.018	.003**	.178	3.237	.075
Zweite Klasse	.259	6.454	.013*	.364	15.159	.000**

Legende: *p<.05 ; **p<.01 ; Stadt: Signifikanz: df 1= 1 , df 2 = 92; Land: Signifikanz: df 1= 1, df 2= 102

Um zu untersuchen ob sich Stadt- oder Landkinder genauer einschätzen, wurde das obige Verfahren mit den „Fisher's-Z“ für unabhängige Stichproben verwendet. Die Differenzen der „Fisher's-Z“ für die drei Jahre und das Insgesamt über alle Jahre bewegen sich zwischen $Z=0.037$ und $Z=0.130$. Mit dem t-Test für abhängige Stichproben wurde die grösste Differenz zwischen Stadt- und Landkindern geprüft. Im ersten Schuljahr, wo sich die grösste Differenz ergab, findet sich ein $t=0.905$, was nicht signifikant ist. Es ergeben sich also zwi-

schen Stadt- und Landkindern auf keiner Stufe signifikante Unterschiede in der Schätzgenauigkeit.

Hypothese 4.4., wonach sich Mädchen genauer einschätzen als Knaben, kann für das Monopedale Überhüpfen nicht bestätigt werden

Hypothese 5.3., wonach städtische Schulklassen eine tiefere Schätzgenauigkeit aufweisen als ländliche, kann für das Monopedale Überhüpfen nicht bestätigt werden.

4.4. Ergebnisse im Sport: Balancieren rückwärts

Im Sport wurden die Einschätzungen der Leistungen und die Leistungen über die beiden Variablen „Balancieren rückwärts“ und „Hüpfen“ erhoben und gemessen. Im Unterschied zur Mathematik konnte die Einschätzung der eigenen Leistung in diesen beiden Aufgaben einerseits auf der Treppenskala vorgenommen werden, andererseits aber auch an der Aufgabe selbst. Im Folgenden werden die Resultate zu „Balancieren rückwärts“ (im folgenden als „Balancieren“ bezeichnet) vorgestellt. Zuerst wieder die Befunde zu den Leistungen, dann die Einschätzung der Leistung, zuerst am Massstab der Treppe, dann am Massstab der Aufgabe selbst. Anschliessend folgen die Ergebnisse zur Über-, respektive Unterschätzung und zur Schätzgenauigkeit. Die gewählten Verfahren sind dieselben wie im Hüpfen.

4.4.1. Die Leistungen in Balancieren rückwärts

Mittels einer dreifaktoriellen ANOVA mit den unabhängigen Faktoren Geschlecht und Milieu („between-subjects“) und der Klassenstufe („within-subjects“) wurden die Leistungen in einem einzigen Schritt berechnet.(Tab.4-4-1).

Tabelle 4-4-1: ANOVA zu den Leistungen in Balancieren rückwärts

Quelle	Df	F	P
Klassenstufe	2	28.879	0.000**
Geschlecht	1	2.971	0.086
Milieu	1	6.160	0.014*
Geschlecht x Milieu	1	0.169	0.681
Stufe x Geschlecht	2	2.543	0.081
Stufe x Milieu	2	0.575	0.562
Stufe x Geschlecht x Milieu	2	0.091	0.911

Es ergeben sich keine signifikanten Interaktionen. Der Klassenstufeneffekt und der Milieueffekt sind signifikant. Die Kontrastanalyse zeigt, dass sich die Durchschnittswerte der zurückgelegten Distanz (KG: 164 cm; 1.Kl.: 133 cm; 2.Kl.: 191 cm) in jedem Fall signifikant unterscheiden (KG-1.Kl.: 32 cm, $p < 0.01$; 1.Kl.-2.Kl.: 58 cm, $p < 0.01$ und KG-2.Kl.: 27 cm, $p < 0.01$). Im Kindergarten betrug die Abweichung von einer mittleren Leistung (Mitte des Balkens) ungefähr eine Fusslänge, in der ersten Klasse war es durchschnittlich ein ganzer Schritt. Die Auf-

gabenschwierigkeit vom Kindergarten zur zweiten Klasse im Balancieren hätte aber gleich bleibend sein sollen, weshalb erwartet wurde, dass die Leistung innerhalb der Zufallswahrscheinlichkeit um den durchschnittlichen Wert von 200 cm schwankt. Dies ist nur in der zweiten Klasse der Fall ($t_{(193)} = 5.189$, $p > 0.10$), im Kindergarten und der ersten Klasse bleiben die Durchschnittsleistungen signifikant unter den Erwartungen (beide $p < 0.01$). Auch aus der Sicht der Kinder entpuppte sich das Balancieren vermutlich als schwierig, insbesondere in der ersten Klasse mit einer Distanz von nur 133 cm anstelle von 200 cm. Diese unterdurchschnittliche Leistungserfahrung im ersten Schuljahr konnte sich aber erst auf die Einschätzungen in der zweiten Klasse auswirken, da die Leistungserbringung nach der Einschätzung erfolgte. Möglich ist aber auch, dass diese Aufgabe bei den Kinder als „schwierig“ galt, weil die maximal mögliche Leistung beim Balancieren augenfällig war: Im Hüpfen und in Mathematik hingegen sahen die Kinder nach ihrer Leistungserbringung nicht, welches die höchste Leistung ihrer Klassenstufe war.

Die Landkinder erbrachten signifikant bessere Leistungen als die Stadtkinder (Tab.4-4-2). Der Geschlechtsunterschied dagegen ist nicht signifikant. Es gibt allenfalls einen Trend ($p = 0.086$), der zeigt, dass die Mädchen im Balancieren leistungsstärker sind als die Knaben (Tab.4-4-2), weshalb die Befunde gleichwohl vorgestellt werden.

Tabelle 4-4-2: Leistungen im Balancieren in cm

	Knaben	Mädchen	Stadt	Land
Kindergarten	155 cm	176 cm	153 cm	177 cm
1.Klasse	133 cm	135 cm	125 cm	142 cm
2.Klasse	173 cm	213 cm	173 cm	210 cm

Kinder erbringen von Klassenstufe zu Klassenstufe in Balancieren, Hüpfen und Mathematik in jedem Fall signifikant höhere Leistungen. In Mathematik drückt sich das in höheren Rohwertpunkten aus, im Hüpfen indem die Kinder höher springen, im Balancieren liegt der Fall aber anders: Das Balancieren wird immer etwas schwieriger, weil ja der Balancierbalken konisch ist, also mit der Breite von 5.5 cm beginnt und mit 2.5 cm endet. Der Balken besteht aus 6 Teilstücken, woraus sich eine Gesamtdistanz von sechs Metern ergibt. Wie in Mathematik und im Hüpfen auch wurde aber eine adaptive Aufgabe gestellt, das heisst im Kindergarten wurden die ersten vier Teilbalken vorgelegt, im ersten Schuljahr die Teilbalken 2 bis 5 und in der zweiten Klasse die Balken 3 bis 6. Jede dieser Aufgaben hätte so für die jeweilige Klassenstufe gleich schwierig sein sollen (gemäss Pilotierungen und KTK-Normierung). In der Auswertung der ANOVA (Tab.4-4-1) hätte das bedeutet, dass der Klassenstufeneffekt beim

Hüpfen nicht hätte signifikant sein dürfen. Die Kinder hätten einfach durchschnittlich die Mitte der vier Meter erreichen sollen. Die vorliegenden Befunde zeigen nun, dass die Aufgabenschwierigkeit in allen Vergleichen signifikant unterschiedlich war. Neben dieser statistischen Ungleichheit haben aber vermutlich die Kinder auch subjektiv gespürt, dass die Aufgabe nicht einfach war, denn schon im Kindergarten blieben sie leistungsmässig unter den Erwartungen. Die Einschätzung im ersten Schuljahr ist vermutlich, sofern sich Kinder am eigenen Lernerfolg messen, bereits etwas von der (statistisch signifikanten) unterdurchschnittlichen Leistung im Kindergarten beeinflusst. Die noch tiefere durchschnittliche Leistung im ersten Schuljahr kann sich dann auf die Einschätzung im zweiten Jahr niederschlagen.

4.4.2 Höhe der Einschätzung auf der Treppenskala

Die Gesamtzahl der Kinder, die alle drei Jahre die Aufgabe bewältigt haben, liegt etwas unterhalb derjenigen in Mathematik, weil einige Kinder über längere Zeit verletzt waren und nicht nachuntersucht werden konnten (Tab.4-4-3).

Tabelle 4-4-3: Lageparameter und Streuungsmasse der Selbsteinschätzung
Im sozialen Vergleich (Treppenstufen) in Balancieren vom Kindergarten zur zweiten Klasse

	Schätzung der Leistung, Kinderg.	Schätzung der Leistung, 1.Klasse	Schätzung der Leistung, 2.Klasse
N	194	194	194
Arithm. Mittel	0.56	0.19	0.18
Median	0	0	0
Modus	0	0	0
Std. Abweichung	2.904	2.691	2.196
Varianz	8.435	7.240	4.822
Minimum	-8	-8	-6
Maximum	8	8	8

(Mittelwert .56 bedeutet eine Abweichung der Einschätzung um etwas mehr als eine halbe Treppenstufe nach oben)

Das arithmetische Mittel liegt in allen Jahren über dem Wert 0. Im Kindergarten beträgt die Abweichung der Einschätzung im t-Test für eine Stichprobe $t(193)=2.670$, $p<0.01$ und ist damit signifikant. In der ersten Klasse beträgt sie $(t(193)=0.961$, $p>0.10$, in der zweiten Klasse $t(193)=1.144$, $p>0.10$ und ist in beiden Fällen nicht signifikant. Die Mittelwerte sind zwar alle positiv abweichend, die Kinder schätzen die eigene Leistung aber nur im Kindergarten signifikant überdurchschnittlich hoch ein. Die durchschnittliche Einschätzhöhe nimmt ab, wobei der Abfall vom Kindergarten zur ersten Klasse signifikant ist und dann faktisch nicht mehr weitergeht.

Die Hypothese 1.1. wird teilweise bestätigt, die Kinder schätzen sich in „Balancieren rückwärts“ durchschnittlich nur im Kindergarten signifikant über dem erwarteten Mittelwert (=0) ein. In der ersten und zweiten Klasse schätzen sie sich auch über dem Mittelwert ein, aber nicht signifikant.

Median und Modus liegen durchwegs bei 0, die grösste Anzahl Einzeleinschätzungen der Leistung findet sich also genau in der Skalenmitte. Sie sinkt zuerst von 64 auf 56 und steigt dann auf 70 (Abb.4-4-1). Über dem Median schätzen sich im Kindergarten 75 (39.5%) der Kinder ein, unterhalb sind es 61 (32.%). In der ersten Klasse liegen 62 (31.3%) der Kinder über dem Median und 82 (42.3%) darunter, während es in der zweiten Klasse 53 (27.3%) sind, die über, respektive 77 (39.7%) die unter dem Median liegen. In den beiden Schuljahren schätzen sich also mehr Kinder unter dem Klassenmittel ein als darüber. Die Befunde sind für den Kindergarten und das erste Schuljahr sehr ähnlich wie diejenigen für das monopedale Überhüpfen. Im Vergleich mit der Mathematik liegt der Median im Balancieren und im Hüpfen in beiden Schuljahren eine Stufe tiefer.

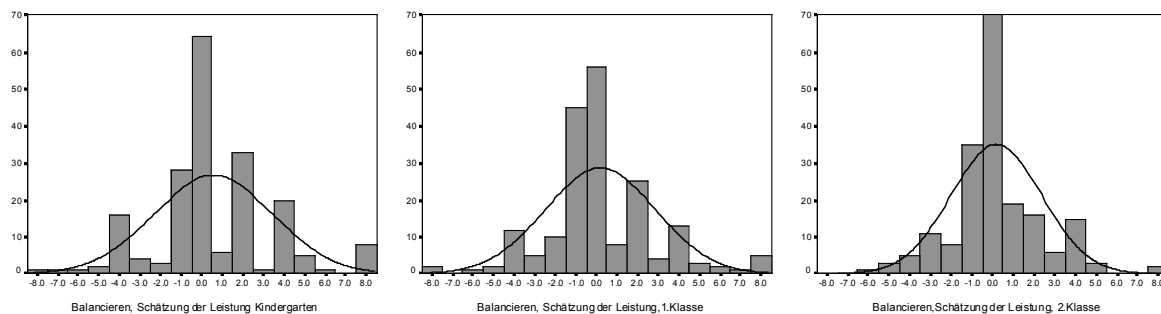


Abbildung 4-4-1: Anzahl der Einschätzungen pro Treppenstufe der eigenen Leistungsfähigkeit in Balancieren rückwärts vom Kindergarten zur zweiten Klasse

Der Vergleich der Standardabweichungen, respektive der Varianzen, ergibt gut übereinstimmende Befunde mit denjenigen beim Hüpfen. Die Varianz im Kindergarten weicht gegenüber derjenigen in der zweiten Klasse signifikant ($F(1.39)=1.749$, $p<0.01$) ab. Ebenfalls signifikant ist der Vergleich der ersten Klasse mit der zweiten Klasse ($F(1.39)=1.501$, $p<0.01$). Der Vergleich des Kindergartens mit der ersten Klasse ist nicht signifikant ($F(1.20)=1.165$, $p>0.10$). Vom Kindergarten zur ersten Klasse kommt es damit im Balancieren, anders als im Hüpfen, noch zu keiner signifikanten Abnahme der Varianz, dies geschieht erst ein Jahr später. Die schwachen Leistungen der Kinder haben sich vielleicht trotzdem in die Verteilung der Einschätzungen eingeschlichen, denn Einschätzungen auf Stufe -3 und tiefer bleiben über alle Jahre häufig, solche oberhalb der Stufe 4 gehen dagegen zurück. Aufgrund

der Symmetrie, der konstanten Gruppengrößen und der grossen Anzahl von Kindern, die gemessen wurden, können die Varianzunterschiede als für die vorliegende Untersuchung unerheblich bezeichnet werden (Bortz, 1993).

4.4.3 Befunde zur Einschätzhöhe unter Berücksichtigung von Klassenstufe, Geschlecht und Milieu

Mittels einer dreifaktoriellen ANOVA mit den unabhängigen Faktoren Geschlecht und Milieu („between-subjects“) und der Klassenstufe („within-subjects“) wurden die Befunde in einem einzigen Schritt berechnet. (Tab.4-4-4).

Tabelle 4-4-4: ANOVA: Einschätzung auf der Treppenskala in Balancieren rückwärts

Quelle	Df	F	P
Klassenstufe	2	1.833	0.161
Geschlecht	1	1.265	0.262
Milieu	1	0.324	0.324
Geschlecht x Milieu	1	0.169	0.681
Stufe x Geschlecht	2	0.814	0.444
Stufe x Milieu	2	0.833	0.428
Stufe x Geschlecht x Milieu	2	2.216	0.115

Kein einziger Effekt ist signifikant. Keine der Klassenstufen unterscheidet sich signifikant. Über die drei Messzeitpunkte verändert sich die Einschätzhöhe nicht signifikant, es ergibt sich allenfalls ein schwacher linearer Trend ($p=0.137$). Im Kindergarten ist die Einschätzhöhe am höchsten, sie nimmt zur ersten Klasse hin ab und wird dann nochmals, wenn auch nur minim, kleiner.

Die Hypothese 2.3., wonach die Einschätzhöhe im Balancieren,

a) im Kindergarten im Vergleich zum ersten Schuljahr gleich hoch ist, wird nicht verworfen

b) in der ersten Klasse im Vergleich zur zweiten Klasse gleich hoch ist, wird nicht verworfen

der Unterschied ist in beiden Fällen nicht signifikant.

Die Hypothese 2.4., wonach die Einschätzhöhe im Kindergarten im Vergleich zur zweiten Klasse gleich hoch ist, kann beibehalten werden. Die Einschätzhöhe im Kindergarten unterscheidet sich nicht signifikant von derjenigen in der zweiten Klasse.

In Tab.4-4-4 ergaben sich weder ein signifikanter Milieu- oder Geschlechtseffekt, noch irgendwelche signifikante Interaktionen zwischen Stufe und Geschlecht oder Stufe und Milieu.

Die Hypothese 4.2., wonach sich die Knaben im Sport höher einschätzen als die Mädchen, kann für das Rückwärtsbalancieren nicht bestätigt werden.

Die Hypothese 5.2., wonach sich die urbanen Klassen im Sport höher einschätzen als die ruralen, kann für das Rückwärtsbalancieren nicht bestätigt werden.

Auch wenn sich keine signifikanten Unterschiede in den beiden unabhängigen Variablen Geschlecht und Milieu ergeben haben, so sind die Befunde mit einer Ausnahme (Milieu, erste Klasse) immer erwartungstreu: Knaben schätzen sich etwas höher ein als Mädchen. Urbane Klassen liegen im Kindergarten leicht, in der zweiten Klasse dann erheblich über den Einschätzungen der ruralen Klassen.

Mit klassenstufenweisen UNIANCOVAs werden nun jahrgangsweise die Einschätzungen auf Geschlechts- und Milieuunterschiede geprüft und zwar unter Konstanthaltung der Leistung. Bei dieser Prüfung zeigt sich, dass die Leistung nur in der zweiten Klasse einen signifikanten Einfluss auf die Einschätzung hat ($F=4.935$, $p<0.05$). Im Kindergarten ist der Einfluss mit $F=0.326$, $p>0.10$ und im ersten Schuljahr mit $F=2.993$, $p=0.085$ dagegen nicht signifikant (Tab.4-4-5).

Tabelle 4-4-5: Klassenstufenweise UNIANCOVA in Balancieren rückwärts mit Geschlecht und Milieu als unabhängigen Faktoren und der Leistung als Kovariate

Quelle	Df	F	P
Kindergarten			
Geschlecht	1	0.028	0.868
Milieu	1	0.347	0.557
Geschlecht x Milieu	1	1.199	0.275
Erste Klasse			
Geschlecht	1	0.300	0.584
Milieu	1	0.289	0.591
Geschlecht x Milieu	1	0.362	0.548
Zweite Klasse			
Geschlecht	1	0.098	0.755
Milieu	1	0.162	0.668
Geschlecht x Milieu	1	1.993	0.159

Die Interaktionen sind nicht signifikant. Die Haupteffekte sind darum direkt diskutierbar. Nur in der zweiten Klasse hat die Leistung einen signifikanten Varianzanteil an der Einschätzung. Milieu und Geschlecht haben keinen signifikanten Effekt. Auch unter Kovarianz der Leistung und klassenstufenweisem Vergleich ergeben sich damit keine Unterschiede in den beiden unabhängigen Variablen.

Angesichts der unterdurchschnittlichen Leistungen der Kinder beim Balancieren (vor allem im ersten Schuljahr) wäre es nicht erstaunlich, wenn die Einschätzung der eigenen Leistung an der Aufgabe über die drei Messzeitpunkte hinweg sinkt: Tatsächlich bewegen sich die Schätzungen im Durchschnitt von 241cm über 221cm auf 204 cm in der zweiten Klasse stetig der Mitte zu, allerdings ergibt sich eine Interaktion von Stufe x Geschlecht (Tab.4-4-6).

Tabelle 4-4-6: Dreifaktorielle ANOVA mit Einschätzung der Leistung an der Aufgabe (in cm) in Balancieren rückwärts als abhängiger Variable, sowie Stufe, Geschlecht und Milieu

Quelle	Df	F	P
Klassenstufe	2	3.992	0.000**
Geschlecht	1	0.540	0.463
Milieu	1	0.650	0.421
Geschlecht x Milieu	1	0.169	0.681
Stufe x Geschlecht	2	3.992	0.021*
Stufe x Milieu	2	0.840	0.427
Stufe x Geschlecht x Milieu	2	0.800	0.444

Die nähere Betrachtung dieser Interaktion (Tab.4-4-7 und Abb.4-4-2) zeigt die Einschätzunterschiede beim Geschlecht:

Tabelle 4-4-7: Einschätzung im Balancieren, an der Aufgabe, getrennt nach Geschlecht und klassenstufenweise UNIANOVAs

Einschätzung	Knaben	Mädchen	DF	F	P
Kindergarten	250 cm	229 cm	1	2.792	0.096
Erste Klasse	215 cm	228 cm	1	1.023	0.313
Zweite Klasse	210 cm	198 cm	1	1.858	0.174

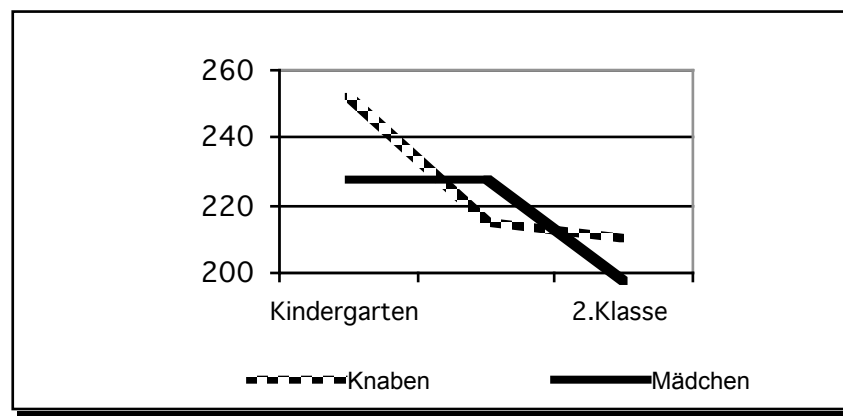


Abbildung 4-4-2: Einschätzung der Distanz im Balancieren rückwärts im Kindergarten, der ersten und zweiten Klasse, unter Kovarianz der Leistung

Aufgrund der disordinalen Interaktion kann der Abfall der Einschätzung weder eindeutig der Klassenstufe noch dem Geschlecht zugeordnet werden. Es können nur die Einfacheffektanalysen aus Tab.4-4-6 interpretiert werden, dort finden sich auf keiner Klassenstufe signifikante Geschlechtsunterschiede. Im Kindergarten und der zweiten Klasse schätzen sich die Knaben leistungsstärker ein, in der ersten Klasse die Mädchen.

Vergleicht man nun die Einschätzung an der Aufgabe mit derjenigen auf der Treppenskala, dann ergeben sich unterschiedliche Befunde: Während bei der Einschätzung an der Treppe keine signifikante Veränderung in der Einschätzhöhe gefunden wurde, nimmt die Einschätzhöhe an der Aufgabe markant ab. Diese signifikante Abnahme wurde nicht erwartet, wie im Hüpfen wurde eine gleich bleibende Einschätzhöhe prognostiziert. Der Grund für diesen erwartungswidrigen Befund bei der Einschätzung der Leistung an der Aufgabe kann nur vermutet werden, zwei Erklärungen stehen im Vordergrund: Erstens könnte sich die Vermutung bestätigen, dass die Kinder Leistungseinschätzungen vorwiegend auf der Basis von eigenen Leistungsfortschritten und Erfolgen vornehmen. Da die Kinder die Aufgabe im Kinder-

garten unter dem erwarteten Leistungsdurchschnitt bewältigten und sie in der ersten Klasse noch schwächer waren, kann dies nur eine sinkende Beurteilung nach sich ziehen. Zweitens konnten die Kinder nur im Balancieren direkt sehen, was die absolute Leistungsspitze ihrer Klassenstufe zu leisten im Stande ist, weil sie den Abstand zum Ende des Balkens sahen. Das erlaubte eine direkte Leistungsrückmeldung. Im Hüpfen wurde zwar bei der Instruktion ebenfalls gezeigt, wie hoch die Leistungsspitze springt, aber dann begannen die Kinder mit dem Hüpfen wesentlich tiefer. Sie dürften bis zur Erreichung ihrer Leistungsgrenze nicht mehr genau gewusst haben, was eine absolute Spitzenleistung wäre.

Dies sind allerdings nur Vermutungen, die Befunde von Wigfield et al. (1997) zeigten ebenfalls eine (allerdings nicht signifikante) Abnahme der Einschätzhöhe vom ersten zum zweiten Schuljahr. Studien die den Kindergarten mit einbeziehen gibt es keine. In Kap. 2.6.1 wurde angenommen, dass sich die Kinder im Sport ab der ersten Klasse sowohl am Leistungsfortschritt, wie auch am sozialen Vergleich orientieren, wenn sie eine Leistungseinschätzung vornehmen. Weil aber in dieser Altersphase die Leistungsfortschritte eher gering sind und bereits im Kindergarten Leistungsvergleiche möglich sind, sei allenfalls mit einem Trend zur Abnahme der Einschätzhöhe zu rechnen.

Fazit: Die klassenstufengerechte Adaption der Aufgabe (in jedem Jahr ist die zu lösende Aufgabe gleich schwierig) ist nicht ganz gelungen, war die Aufgabe doch in der ersten Klasse auch subjektiv eindeutig schwerer als in der zweiten Klasse und im Kindergarten.

Es könnte also sein, dass die Abnahme der Einschätzhöhe auf die schwierigere Aufgabenstellung sowohl im Kindergarten, wie in der ersten Klasse zurückgeht. Das kann überprüft werden, indem die Daten aus der Kontrollgruppenuntersuchung für die Konstruktion einer Querschnittuntersuchung genutzt werden. Dabei lassen sich die Kinder der Kindergartenstichprobe (N=194) mit denjenigen der Kontrollgruppe 1 (N=101) und der Kontrollgruppe 2 (N=96) vergleichen. Der Vorteil der querschnittlichen Betrachtungsweise liegt darin, dass so nur Kinder verglichen werden, die ihre Einschätzungen erstmalig machen, also vor dieser Einschätzung die Aufgaben noch nie lösten. In Tab.4-4-8 sind die Befunde der beiden Untersuchungsarten dargestellt.

Tabelle 4-4-8: Vergleich der Einschätzung im Balancieren zwischen der Längs- und der Querschnittmessung

	Kindergarten	Erste Klasse	Zweite Klasse
Längsschnitt	241 cm	221 cm	204 cm
Querschnitt	241 cm	224 cm	210 cm

Aus den Kontrollgruppenuntersuchungen ging hervor, dass sich die Einschätzungen zwischen den Erst- und Zweitklässlern aus der Längsschnittuntersuchung nicht signifikant von denjenigen aus den Kontrollgruppen unterscheiden. Die Einschätzunterschiede bei querschnittlicher Betrachtung zeigen nun vom Kindergarten zur ersten Klasse einen signifikanten Abfall (17cm, $p < 0.05$) und von der ersten zur zweiten Klasse ebenfalls eine signifikant tiefere durchschnittliche Einschätzhöhe (14cm, $p < 0.05$).

Fazit: Aufgrund der disordinalen Interaktion Stufe x Geschlecht lässt sich nicht eindeutig belegen, ob die Abnahme der Einschätzhöhe an der Aufgabe auf den Klassenstufen- oder auf den Geschlechtseffekt zurückgeht. Die unterschiedliche Aufgabenschwierigkeit in den einzelnen Klassenstufen kann das Absinken der Einschätzhöhe auch nicht erklären, denn die querschnittlichen Befunde zeigen ebenfalls eine signifikante Abnahme der Einschätzhöhe.

Auch beim monopetalen Überhüpfen war die Einschätzhöhe, gemessen an der Aufgabe, über die Jahre durchschnittlich sinkend, allerdings nicht signifikant. In den beiden Sportaufgaben, die eine Einschätzung der eigenen Leistung direkt an der Aufgabe zulassen, werden damit die Befunde von Wigfield et al. (1997) bestätigt.

Die Einschätzungen der Leistung auf der Treppenskala im Balancieren und im Monopetalen Überhüpfen unterscheiden sich im Kindergarten nicht signifikant von denen in der zweiten Klasse. Dies entspricht den Erwartungen und zeigt, dass sich Unterschiede ergeben, ob nun eine am sozialen Massstab oder eine an der Aufgabenschwierigkeit fokussierte Einschätzung vorgenommen wird. Das stützt die Vermutung aus Kap.2.6, wonach die soziale Bezugsnorm im Sport schon im ersten Schuljahr bedeutsam wird insofern, als sich die Kinder offenbar bei der Einschätzung an der Treppe in der Balancieraufgabe von der Aufgabenschwierigkeit wenig beeindruckt liessen und sich nicht signifikant (wenn auch tendenziell) tiefer einschätzten, wie dies bei der Einschätzung an der Aufgabe der Fall war.

4.4.4 Über- oder Unterschätzung als Differenz zwischen Einschätzung und Leistung im Balancieren rückwärts

Um Klarheit darüber zu gewinnen, ob sich die Kinder unter- oder überschätzen und ob es Geschlechts- und Milieuunterschiede gibt, wird von der Einschätzung der Leistung (an der Aufgabe) die erbrachte Leistung direkt in Abzug gebracht (Tab.4-4-9).

Tabelle 4-4-9: Dreifaktorielle ANOVA mit Über-/ Unterschätzung als abhängiger Variable

Quelle	df	F	P
Klassenstufe	2	6.343	0.002
Geschlecht	1	4.728	0.031
Milieu	1	8.566	0.004
Geschlecht x Milieu	1	0.452	0.502
Stufe x Geschlecht	2	5.261	0.006
Stufe x Milieu	2	0.744	0.475
Stufe x Geschlecht x Milieu	2	0.594	0.551

Eine signifikante Interaktion besteht zwischen Geschlecht und Stufe, sodass die Klassenstufen einzeln zu betrachten sind (Tab.4-4-10):

Tabelle 4-4-10: Mittelwerte und UNIANOVAs der Überschätzung im Balancieren rückwärts

Jahr	Mittelwert Knaben	Mittelwert Mädchen	df	F	P
Kindergarten	62.94 cm	17.78 cm	1	6.922	0.009
1. Klasse	16.07 cm	27.01 cm	1	0.525	0.470
2.Klasse	29.10 cm	-22.52 cm	1	11.728	0.001

Signifikant sind die Mittelwertunterschiede im Kindergarten, die Knaben überschätzen sich mehr als die Mädchen. In der ersten Klasse überschätzen sich die Mädchen mehr als die Knaben, aber nicht signifikant. In der zweiten Klasse unterschätzen sich die Mädchen, so dass der Mittelwertunterschied zwischen den Geschlechtern signifikant wird.

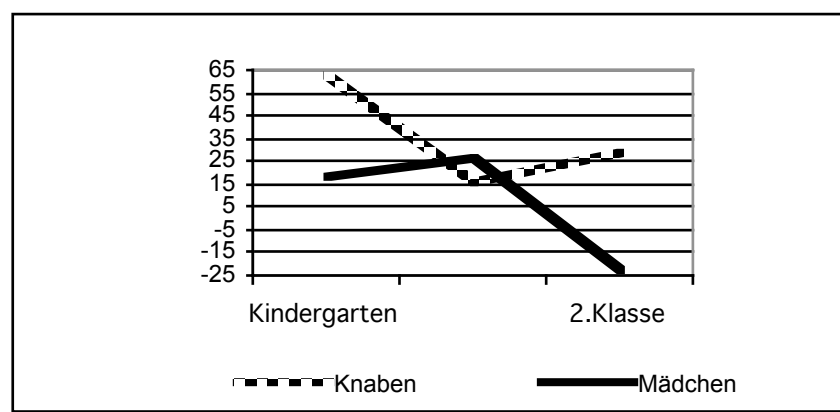


Abbildung 4-4-3: Unter- / Überschätzung der Leistung in cm, im Balancieren, im Kindergarten, der ersten und zweiten Klasse

Die Entwicklung der Überschätzung verläuft bei den Mädchen signifikant anders als bei den Knaben: Zuerst nimmt die Überschätzung mässig zu und dann so stark ab, dass es zu einer Unterschätzung kommt. Bei den Knaben nimmt die Überschätzung zuerst markant ab um dann ganz wenig zuzunehmen. Aufgrund dieser Unterschiede können die Geschlechter nur getrennt betrachtet werden, nun zuerst zu den Knaben (Tab.4-4-11).

Tabelle 4-4-11: Zweifaktorielle ANOVA mit Über-/ Unterschätzung der Knaben als abhängiger Variable und Milieu als unabhängiger Variablen.

Quelle	Df	F	P
Klassenstufe	2	10.85	0.000***
Milieu	1	0.854	0.358
Stufe x Milieu	2	0.921	0.400

Die Überschätzung nimmt signifikant ab, wie der Klassenstufeneffekt zeigt. Die post hoc-Kontrastanalyse zeigt, dass die Abnahme der Überschätzung bei den Knaben vom Kindergarten zur ersten oder zweiten Klasse signifikant ist (KG-1.Kl.: 49cm, $p < 0.01$; KG-2.Kl.: 35cm, $p < 0.05$). Die Zunahme der Überschätzung vom ersten zum zweiten Schuljahr ist dagegen nicht signifikant (1.Kl.-2.Kl.: 13cm, $p > 0.1$). Die Regressionsanalyse ergibt einen signifikanten linearen ($F=6.71$, $p < 0.05$) und einen signifikanten quadratischen ($F=6.52$, $p < 0.05$) Effekt. Die Überschätzung nimmt also bei den Knaben signifikant diskontinuierlich ab. Die Knaben überschätzen sich im Kindergarten und der zweiten Klasse signifikant (KG: $t(99)=5.475$, $p < 0.001$; 2.Kl.: $t(99)=2.986$, $p < 0.01$).

Hypothese 1.2. wird für die Knaben mehrheitlich bestätigt: Im Kindergarten und der zweiten Klasse überschätzen sich die Knaben im Balancieren rückwärts signifikant. In der ersten Klasse überschätzen sie sich ebenfalls, aber nicht signifikant.

Die Überschätzung im Rückwärts Balancieren nimmt bei den Knaben vom Kindergarten zur zweiten Klasse signifikant ab.

Nun folgen die Befunde für die Mädchen (Tab.4-4-12).

Tabelle 4-4-12: Zweifaktorielle ANOVA mit Über-/ Unterschätzung der Mädchen als abhängiger Variable und Milieu als unabhängiger Variablen.

Quelle	Df	F	P
Klassenstufe	2	5.100	0.007**
Milieu	1	10.100	0.002**
Stufe x Milieu	2	0.460	0.632

Die Überschätzung nimmt bei den Mädchen signifikant ab, wie der Klassenstufeneffekt zeigt. Die post hoc-Kontrastanalyse zeigt, dass die Überschätzung bei den Mädchen vom Kindergarten zur ersten Klasse zunimmt, aber nicht signifikant. Von der ersten zur zweiten Klasse nimmt die Überschätzung signifikant ab (1.Kl.-2.Kl.: 46cm, $p < 0.01$) und mündet in eine Unterschätzung. Vom Kindergarten zur zweiten Klasse ist die Abnahme nicht signifikant, es ergibt sich allenfalls ein Trend zur Abnahme (KG-2.Kl.: 38cm, $p = 0.077$). Die Regressionsanalyse ergibt einen signifikanten linearen ($F = 5.13$, $p < 0.05$) und einen signifikanten quadratischen ($F = 5.05$, $p < 0.05$) Effekt. Die Überschätzung nimmt also bei den Mädchen signifikant diskontinuierlich ab: Im Kindergarten und der ersten Klasse überschätzen sich die Mädchen, um sich dann in der zweiten Klasse zu unterschätzen.

Die Mädchen überschätzen sich im Kindergarten nicht signifikant. In der ersten Klasse hingegen überschätzen sie sich signifikant (1.Kl.: $t_{(93)} = 2.598$, $p < 0.05$), die Unterschätzung in der zweiten Klasse ist dagegen knapp nicht signifikant (2.Kl.: $t_{(93)} = 1.944$, $p = 0.055$).

Hypothese 1.2. wird für die Mädchen nicht bestätigt: Im Kindergarten überschätzen sie sich im Balancieren rückwärts nicht signifikant. In der ersten Klasse überschätzen sie sich signifikant, während sie sich in der zweiten Klasse knapp nicht signifikant unterschätzen.

Die Überschätzung im Balancieren rückwärts nimmt bei den Mädchen vom Kindergarten zur zweiten Klasse nicht signifikant ab.

Der Milieueffekt ist signifikant. Die Unterschiede zwischen Stadt- und Landkindern nehmen über die Jahre noch zu, aber nicht so stark, dass daraus eine Interaktion mit der Stufe resultiert. Die Überschätzung ist bei den Landkindern nach zwei Jahren negativ, diese Kinder unterschätzen sich also, weshalb hier die Verläufe trotz fehlender Interaktion mit der Stufe dargestellt werden (Abb. 4-4-4):

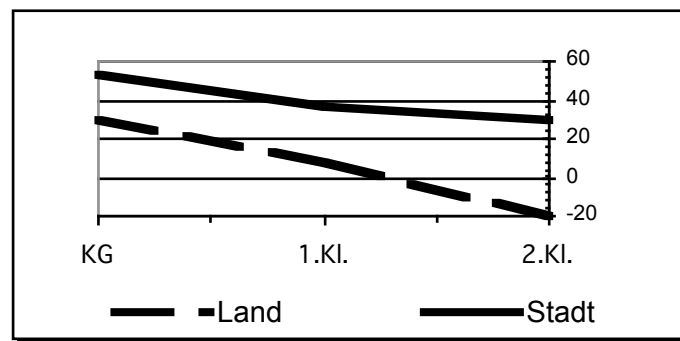


Abbildung 4-4-4: Überschätzung im Balancieren rückwärts, in cm, getrennt nach Land- und Stadtkindern

Hypothese 4.3, wonach sich Knaben mehr überschätzen als Mädchen kann wegen Interaktionseffekten nur klassenstufenweise geprüft werden. Im Kindergarten und der zweiten Klasse überschätzen sich die Knaben signifikant stärker als die Mädchen. In der ersten Klasse überschätzen sich die Mädchen stärker als die Knaben, allerdings nicht signifikant.

Hypothese 5.2. wird bestätigt, urbane Kinder überschätzen sich signifikant stärker als rurale.

4.4.5 Überprüfung der Zusammenhangshypothesen

Mit der Überprüfung von Zusammenhangshypothesen gilt es herauszufinden, ob sich die Genauigkeit der Leistungsschätzung im Laufe der Schuljahre signifikant verändert und ob sich Knaben oder Mädchen, Stadt- oder Landkinder signifikant unterschiedlich einschätzen. Die Korrelation von Einschätzung x Leistung überprüft, wie eng die Schätzungen der Kinder mit der tatsächlichen Leistungsfähigkeit übereinstimmen. Beim Hüpfen ist die Einschätzung an zwei Massstäben erfolgt, an der Aufgabe direkt und auf der Treppe (Tab.4-4-13).

Tabelle 4-4-13: Korrelationen von Einschätzung (an der Gruppe und an der Aufgabe) mit der Leistung

	EG1	EG2	EAK	EA1	EA2	LKG	L1K	L2K
Einschätz g.Grp.								
KG	-.011	.040	.199**	.094	.095	.037	.062	.072
1.Kl.		.177*	.257**	.252**	.220**	.111	.131	.042
2.Kl.			.147*	.085	.335**	.169*	.114	.193**
Einschätzg. Aufg.								
KG				.302**	.222**	.140	.181*	.124
1.Kl.					.360**	.156*	.132	.154*
2.Kl.						.205	.220**	.182*
Leistung								
KG.							.345**	.358**
1.Kl.								.282**

(Bedeutung der Abkürzungen, z.B.: EG1=Einschätzung an der Gruppe, respektive auf der Treppe, in der 1.Klasse; EAK=Einschätzung an der Aufgabe im Kindergarten; L2K=Leistung in der 2.Klasse)

Die Einschätzungen auf der Treppe zeigen nur in der zweiten Klasse einen schwachen signifikanten Effekt: Im Vergleich ist damit die Schätzgenauigkeit in dieser Sportaufgabe ähnlich tief ausgefallen, wie im monopodalen Hüpfen, aber tiefer als in Mathematik. Besonders im Kindergarten konnte ein anderes Ergebnis erwartet werden, weil dort ja die Erfahrungs- und Vergleichsmöglichkeiten potentiell besser wären als in Mathematik, wo gar kein entsprechender Fachunterricht möglich ist.

Die Korrelationen (Einschätzung an der Aufgabe x Leistung) sind, mit Ausnahme vom Kindergarten, ähnlich. In der zweiten Klasse ergibt sich ein schwacher Korrelationseffekt. Um die Korrelationen miteinander auf signifikante Unterschiede zu prüfen, wurden die Koeffizienten in „Fishers-Z“ umgerechnet. Anschliessend erfolgte für die klassenstufenweisen Vergleiche eine Signifikanzprüfung für abhängige Stichproben (nach Bortz, 1993, 205). Die Korrelationsdifferenz KG-1.Kl. beträgt 0.008. Diese Abnahme ist nicht signifikant. Die Korrela-

tionsdifferenz zwischen den beiden Schulklassen ist auch klein (0.050) und damit nicht signifikant. Zwischen Kindergarten und zweiter Klasse ist sie noch kleiner.

Die Korrelationen (Einschätzung auf der Treppe x Leistung) zeigen nur in der zweiten Klasse einen schwachen signifikanten Effekt. Sie nehmen von Jahr zu Jahr zu. Die Schätzgenauigkeit ist in der zweiten Klasse im Vergleich zum Kindergarten knapp nicht signifikant höher (0.156; $z=1.5956$, $p=0.055$), alle anderen Differenzen sind tiefer.

Hypothese 3.3., wonach die Schätzgenauigkeit in Balancieren rückwärts von Jahr zu Jahr zunimmt, kann nicht bestätigt werden: Vom Kindergarten zur ersten Klasse nimmt sie (nicht signifikant) zu, vom ersten zum zweiten Schuljahr nimmt sie auch zu (aber eben nicht signifikant).

Hypothese 3.4., wonach die Schätzgenauigkeit in Balancieren rückwärts vom Kindergarten zur zweiten Klasse signifikant zunimmt, kann knapp nicht bestätigt werden.

In Tab.4-4-14 werden die Schätzgenauigkeiten von Knaben und Mädchen auf Signifikanz geprüft. Jahrweise Vergleiche nach dem obigen Verfahren mit den „Fishers-Z“ für unabhängige Stichproben ergeben, dass sich im Kindergarten und der zweiten Klasse die Knaben genauer einschätzen als Mädchen. Am grössten ist die Differenz in der zweiten Klasse. Mit einem z -Wert=1.852 ($p < 0.05$) im t-Test ist der Unterschied signifikant. Die anderen Differenzen sind kleiner und nicht signifikant. Die Schätzgenauigkeit der beiden Geschlechter unterscheidet sich also im Kindergarten und der ersten Klasse nicht signifikant.

Tabelle 4-4-14: Vergleich der Korrelationen (Einschätzung an der Aufgabe x Leistung) von Knaben und Mädchen

	Mädchen			Knaben		
	Korrelation	F-Wert	Signifikanz (zweiseitig)	Korrelation	F-wert	Signifikanz (zweiseitig)
Kindergarten	.148	2.066	.154	.168	2.836	.095
Erste Klasse	.195	3.618	.060	.073	0.520	.472
Zweite Klasse	.071	0.462	.499	.325	11.553	.001*

Legende: * $p < .05$; ** $p < .01$; Mädchen: Sign. bei: $df 1 = 1$; $df 2 = 92$; Knaben: Sign bei $df 1 = 1$; $df 2 = 98$

Bei den Landkindern nimmt die Schätzgenauigkeit laufend zu und erreicht in der zweiten Klasse einen signifikanten Wert bei einer schwachen Korrelation. Dieser Befund ist demjenigen im monopodalen Überhüpfen sehr ähnlich. Die Stadtkinder schätzen sich im Kindergarten und genauer ein. In der ersten Klasse ist es umgekehrt und in der zweiten ist die

Schätzgenauigkeit von Stadt- und Landkindern fast gleich. Es ergeben sich in diesem Jahr schwache signifikante Korrelationen (Tab. 4-4-15).

Tabelle 4-4-15: Korrelationen (Einschätzung x Leistung), getrennt nach Stadt- und Landkindern

	Stadt		Signifikanz	Land		Signifikanz
	Korrelation	F-Wert	(zweiseitig)	Korrelation	F-wert	(zweiseitig)
Kindergarten	.253	6.242	.014*	.061	0.365	.547
Erste Klasse	.080	0.589	.445	.188	3.631	.060
Zweite Klasse	.208	4.133	.045*	.201	4.168	.044*

Legende: * $p < .05$; ** $p < .01$; Stadt: Signifikanz: df 1= 1 , df 2 = 91; Land: Signifikanz: df 1= 1, df 2= 99

Um zu untersuchen ob sich Stadt- oder Landkinder genauer einschätzen, wurde das obige Verfahren mit den „Fishers-Z“ für unabhängige Stichproben verwendet. Der t-Test für die Kindergartenstufe ergab ein $z=1,337$ ($p < 0.10$), die Stadtkinder schätzen sich also nicht signifikant genauer ein. Die anderen Differenzen zwischen Stadt- und Landkindern sind noch kleiner. Zwischen Stadt- und Landkindern ergeben sich in der Schätzgenauigkeit auf keiner Klassenstufe signifikante Unterschiede.

Hypothese 4.4., wonach sich Mädchen genauer einschätzen als Knaben, kann für das Balancieren rückwärts nicht bestätigt werden

Hypothese 5.3., wonach städtische Schulklassen eine tiefere Schätzgenauigkeit aufweisen als ländliche, kann für das Balancieren rückwärts nicht bestätigt werden.

4.5 Schlussbetrachtung und Ausblick

4.5.1 Höhe und Verlauf der Selbsteinschätzung

Die Einschätzhöhe weicht im Kindergarten in Mathematik und den beiden Sportaufgaben signifikant vom erwarteten durchschnittlichen Wert ab. Im ersten und zweiten Schuljahr liegt die Einschätzhöhe der Kinder in Mathematik und im Hüpfen ebenfalls signifikant über dem erwarteten Mittelwert. Im Balancieren liegt die Einschätzhöhe ebenfalls über dem erwarteten Mittelwert, aber nicht signifikant. Die Befunde dieser Arbeit bestätigen damit die Hypothesen, wie auch frühere Befunde (vgl. Kap.2.4). Im Kindergarten und den ersten beiden Schuljahren schätzen Kinder ihre Leistungen über dem erwarteten Durchschnitt ein (Abb.4-5-1).

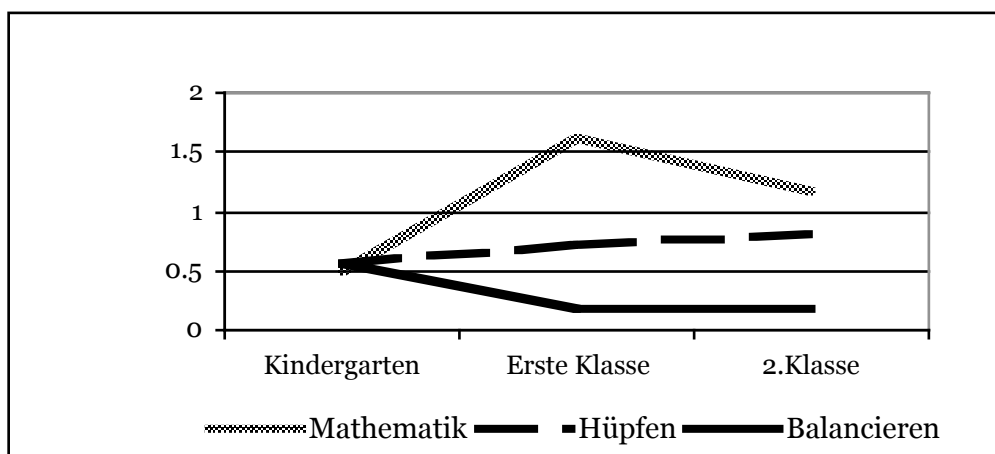


Abbildung 4-5-1: Durchschnittliche Abweichung der Einschätzung vom erwarteten Mittelwert (=0), Skala -8 bis +8 (Treppenstufenskala)

Der Verlauf der Einschätzhöhe (Abb.4-5-1) bestätigt die Hypothesen: In Mathematik nimmt die Einschätzhöhe signifikant zu, während im Sport die Nullhypothesen nicht verworfen werden, denn die Einschätzhöhe ist gleich bleibend.

In Kap.2.3 hatten wir vermutet, dass sich die Kinder mit dem beginnenden Mathematikunterricht im ersten Schuljahr höher einschätzen werden, weil sie sich vorwiegend am individuellen Leistungsfortschritt messen. Kinder des ersten Schuljahres fragen sich, ob sie viel gelernt haben. Für sie gilt, wer viel lernt ist auch gut. In bisherigen Befunden wurde häufig argumentiert, dass die positiven Rückmeldungen der Lehrkraft der Grund für die zunehmende Höhe der Leistungseinschätzung seien. Helmke (1991) meint, dass dies auf Grund der Bezugsnormorientierung der Kinder der Schuleingangsphase (am eigenen Fortschritt)

eher unwahrscheinlich sei. Vielleicht sei es der Wechsel der Lehrperson im Übergang zur Schule und dann vom zweiten zum dritten Schuljahr, der die Einschätzung nach oben verändere. Der Anstieg lässt sich mit dieser These jedoch kaum erklären, wie zwei Beispiele zeigen: Der Wechsel der Lehrkraft führt in den beiden Sportaufgaben nicht zu signifikant anderen Einschätzhöhen und auch die Befunde von Ingold (2002) und Demény (2002) sprechen dagegen. Die Untersuchungen von Nicholls (1978), die in Tab.2-3-1 dargestellt sind, zeigen, dass die Anstrengung, respektive der Lernaufwand, den Kinder betreiben, im Schuleingangsalter von den Kindern als Hauptursache für Erfolg / Misserfolg angesehen wird. Unsere Daten stärken nun die Annahme, dass vorwiegend der beginnende Unterricht in Mathematik die Ursache für die steigende Selbsteinschätzung ist. Kinder befinden sich im Schuleingangsalter auf Stufe 6 der Skala nach Flammer (1996, vgl. Abb.2-2-2). Erst im dritten Schuljahr kommen sie in die nächste Stufe und realisieren, dass neben der Anstrengung auch die Fähigkeiten des Einzelnen den Lernerfolg mitbeeinflussen. Im Laufe des vierten Schuljahres erreichen die Kinder bereits die Stufe 8 nach Flammer (1996) und verfügen nun über ein voll ausgebildetes Attributionsmodell. Sie sind nun kognitiv in der Lage sich realistischer einzuschätzen. Die Befunde von Demény (2002) zum Leistungsbereich Mathematik bestätigen diese Annahmen: Vom dritten zum vierten Schuljahr geht die Einschätzhöhe massiv zurück, nämlich um 1.05 Treppenstufen. Verbindet man nun die Einschätzabweichungen aus der Untersuchung von Demény mit denjenigen aus der vorliegenden Untersuchung, ergibt sich (Abb.4-5-2):

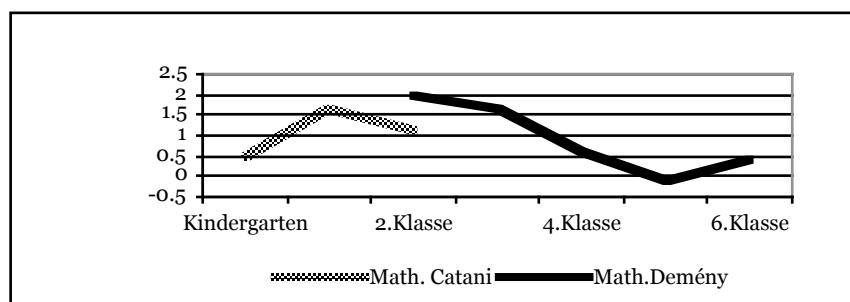


Abbildung 4-5-2: Abweichungen vom Mittelwert (in Treppenstufen) in Mathematik vom Kindergarten bis zum sechsten Schuljahr

Die Befunde bestätigen damit auch die Untersuchungen von Helmke. Zusammenfassend könnte man nun den Verlauf der Einschätzhöhe in etwa so begründen: Vom Kindergarten bis zur zweiten Klasse glauben die Kinder, dass die Anstrengung den Leistungserfolg erklärt (Stufe 6 nach Flammer). Weil sie in Mathematik von der ersten Klasse an diesbezüglich bedeutend mehr investieren als früher, machen die Kinder grosse Fortschritte. Diesen Leistungsfortschritt nehmen sie zum Massstab, um sich in der ersten Klasse dann durchschnittlich (also auf Treppenstufe 1.6) ungefähr so einzuschätzen: „Ich habe sehr viel gelernt

in Mathematik und deshalb bin ich jetzt recht gut bis gut, besser als viele Erstklässler“. In den folgenden Schuljahren werden die Leistungszuwächse der Kinder kleiner (gemäss KAB-C-Normen), die Einschätzung bleibt hoch, geht aber etwas zurück. Im dritten und vierten Schuljahr entwickeln sich die Kinder hin zu Stufe 8 (nach Flammer, 1986). Sie bemerken nun, dass die Leistungen neben der Anstrengung offensichtlich auch von der Begabung abhängig sind. Das lässt das Selbstbild realistischer werden, die Einschätzhöhe sinkt vom zweiten zum dritten und vom dritten zum vierten Schuljahr signifikant ab. Die Entwicklung und Ausdifferenzierung des Attributionskonzeptes ist Ende viertes Schuljahr abgeschlossen. Die Leistungen sinken in der Folge in der Einschätzung bis zum Ende der gemeinsamen Grundschuljahre nicht mehr signifikant. Auch am Ende des sechsten Schuljahres schätzen sich die Kinder signifikant über dem Mittel ein (gemäss Demény, 2002).

Im Sport lassen sich die Einschätzungen zum Balancieren und Hüpfen mit dem Weitsprung in der Untersuchung von Ingold (2002) ergänzen: Die Abweichung vom Mittelwert nimmt vom dritten zum vierten Schuljahr signifikant ab und wird dann nur noch wenig tiefer. Im ersten, zweiten, fünften und sechsten Schuljahr weicht die Einschätzhöhe nicht signifikant vom erwarteten Mittelwert ab, die Werte unterscheiden sich nicht signifikant voneinander (Abb.4-5-3).

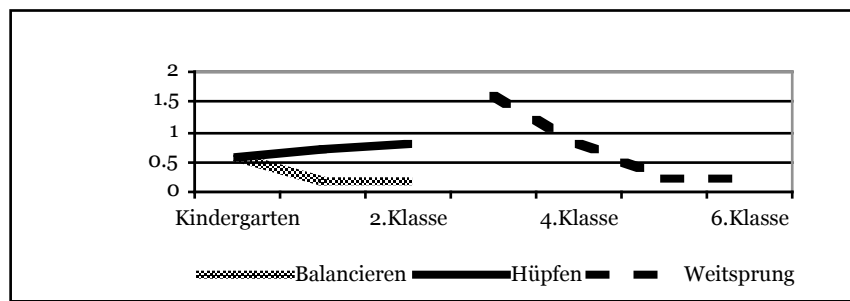


Abbildung 4-5-3: Abweichungen der Einschätzhöhe vom erwarteten Mittelwert (=0) im Sport (Balancieren, hüpfen, Weitsprung) in Treppenstufen.

Die Schätzgenauigkeit nimmt über den Messzeitraum in Mathematik nicht signifikant zu. Erstaunlich ist insbesondere, dass die Schätzgenauigkeit im Kindergarten höher war als im ersten Schuljahr. Dieser Befund ist schwierig zu deuten, denn Kinder im Kindergartenalter haben ja viel weniger Gelegenheit, sich im Rechnen / in Mathematik zu messen, als dann in der Schule. Immerhin liegt die Schätzgenauigkeit im zweiten Schuljahr dann wieder über derjenigen im Kindergarten (aber nicht signifikant). Im Vergleich zu anderen Studien ist sie im zweiten Schuljahr hoch ($r=0.35$). Bei Helmke (1991) ergab sich ein Wert von $r=0.32$, bei Demény (2002) von $r=0.27$, in den anderen Untersuchungen waren die Korrelationen tiefer.

4.5.2 Geschlechts- und Milieudifferenzen

In den vorliegenden Befunden finden sich im Sport und in Mathematik keine signifikanten Leistungsdifferenzen zwischen Knaben und Mädchen. Knaben schätzen sich in Mathematik und im Hüpfen signifikant höher ein als Mädchen. Im Balancieren sind die Einschätzungsdifferenzen nicht signifikant. Kontrolliert man hingegen die Leistung, sieht es anders aus: Auf der Treppenstufenskala ergeben sich nun keine signifikanten Geschlechtsunterschiede mehr. Bei der Einschätzung an der Aufgabe selbst haben sich die Knaben nur beim Hüpfen signifikant leistungsfähiger eingeschätzt als die Mädchen. Dass sich Knaben mehr überschätzen als Mädchen kann durch die vorliegenden Befunde nicht bestätigt werden. Einzig im Balancieren ergaben sich signifikante Unterschiede: Im Kindergarten und der zweiten Klasse überschätzten sich die Knaben signifikant mehr als die Mädchen, während in der ersten Klasse sich die Mädchen, allerdings nicht signifikant, stärker überschätzten.

Die diesbezügliche Befundlage ist für den Schuleingangsbereich unterschiedlich und zudem gibt es sehr wenige Studien zu dieser Frage, wie wir in Kap. 2.5 gesehen haben. Die Hypothesen, wonach sich Knaben höher einschätzen als Mädchen können nur teilweise bestätigt werden und das auch nur, wenn die Leistung nicht kontrolliert wird. Es zeigt sich also, dass die Messung der Leistung bereits im Schuleingangsbereich wichtig ist, will man herausfinden, ob es geschlechtsspezifische Unterschiede in der Einschätzung der eigenen Leistung gibt. Gerade im Kindergartenalter und den ersten beiden Schuljahren wurde in bisherigen Befunden die Leistung häufig durch die Urteile der Erziehenden, respektive der Lehrkräfte, operationalisiert (beispielsweise bei Helmke). In diese Urteile können geschlechtsspezifische Erwartungen einfließen (Tiedemann & Faber, 1994), was die Objektivität beeinträchtigt. Für die vorliegenden Untersuchungen wurde deshalb die Leistung direkt gemessen. Die Entwicklung der Einschätzunterschiede bei Knaben und Mädchen sehen folgendermassen aus: In Mathematik ist der Unterschied im Kindergarten mit 0.29 Treppenstufen noch klein. In der ersten Klasse sind es 0.76, in der zweiten Klasse 0.63 Treppenstufen, um die sich Knaben höher einschätzen. Im Hüpfen bleiben die Einschätzungsdifferenzen stabil, sie betragen im Kindergarten 0.70 Treppenstufen und nehmen auf 0.77 im zweiten Schuljahr zu.

Im Balancieren ergibt sich eine signifikante Interaktion Geschlecht x Klassenstufe. Im ersten Schuljahr sind es die Mädchen, auf den beiden anderen Klassenstufen die Knaben, die sich höher einschätzen. Der Unterschied ist im zweiten Schuljahr kleiner als im Kindergarten. Insgesamt kann kein eindeutiger Trend abgelesen werden, im Sport bleiben die Unterschiede stabil, in Mathematik schätzen sich die Knaben in der Schule signifikant höher ein als die Mädchen.

Unter Kovariation der Leistung können die Hypothesen, wonach sich Knaben leistungsmässig höher einschätzen, nur für das Hüpfen bestätigt werden. Für Mathematik bes-

tätigen sich damit die Befunde von Valeski & Stipek (2001) und Marsh. et al. (1984), wonach sich in der Schuleingangsstufe noch keine signifikanten Einschätzunterschiede fanden. Cole et al. (1998) hat jede Schulstufe vom ersten bis zum sechsten Schuljahr untersucht. Vom dritten bis zum sechsten Schuljahr schätzten sich die Knaben durchgängig höher ein als die Mädchen, ab vierter Klasse signifikant: Die z-standardisierten Residuen der Regression (Einschätzung x Leistung) unserer Untersuchung lauten 0.01 für das erste, 0.00 für das zweite Schuljahr. Bei Cole et al. (1998) sind es dann im dritten 0.08, im vierten 0.11* und im fünften Schuljahr 0.14*. Die Unterschiede nehmen also von Jahr zu Jahr zu.

In Bezug auf die Schätzgenauigkeit lassen sich in allen drei Aufgaben keine signifikanten Geschlechtsunterschiede für die Schuleingangsstufe finden. Dieser Befund stimmt im Sport (Kindergartenstufe) mit demjenigen von Althaus & Bühler (1999) überein, die nur einen Trend zur genaueren Einschätzung bei den Mädchen fanden. In der zweiten Klasse fand Ingold (2002) aber bereits einen signifikanten Unterschied in der Schätzgenauigkeit. Ein signifikanter Unterschied findet sich auch in den vorliegenden Befunden im Balancieren, allerdings nur im zweiten Schuljahr. Im Hüpfen dagegen fand sich keine signifikante Differenz der Schätzgenauigkeit.

Nun noch zu vermuteten Milieudifferenzen. Im Balancieren fand sich ein signifikanter Unterschied, die städtischen Kinder überschätzten sich in allen Jahren signifikant. Dieses Ergebnis ist auf die signifikant höheren Leistungen der ländlichen Kinder zurückzuführen, denn in der Einschätzhöhe (ohne Kontrolle der Leistung) waren die Unterschiede nicht signifikant. Sonst finden sich keine signifikanten Befunde, so sind die Einschätzungen im Hüpfen zwischen Stadt- und Landkindern fast identisch. In Mathematik schätzen sich die Landkinder in den beiden Schuljahren (nicht signifikant) höher ein, im Kindergarten sind es die städtischen Kinder, die sich höher einschätzen.

Gesamthaft betrachtet hat die Kontrolle des Milieus nur einen sehr geringen Einfluss auf die Einschätzungen gehabt. Trotzdem sei ein Ergebnis noch kurz kommentiert: Im Hüpfen, also im Kraftbereich, ergeben sich keine Milieuunterschiede. Hingegen dort, wo es auf die Körperkoordinationsfähigkeiten ankommt, erbrachten die Landkinder signifikant höhere Leistungen als die urbanen Kinder. Die plausibelste Erklärung für dieses Ergebnis kann darin gesehen werden, dass sich die ländlichen Kinder mehr und regelmässiger bewegen, als die städtischen.

4.5.3 Methodisches Vorgehen und Instrumente

Die Datenerhebung erwies sich, wie vermutet, als sehr zeitaufwändig. Die persönliche Betreuung während der gesamten Erhebung (ein Student arbeitet mit einem einzelnen Kind) erwies sich aber als sehr sinnvoll. Die Kinder konzentrierten sich und waren über den gesamten Prüfzeitraum engagiert. Durch die persönliche Betreuung gelang es auch, allfällige Missverständnisse in der Instruktion zu klären. Die Einschätztreppe ist zwar ein anschauliches Instrument, damit die Kinder aber diejenige Treppenstufe bezeichnen, die sie auch wirklich meinen, ist die Kontrolle der Antwort sehr wichtig. Günstig ist die Einschätztreppe auch in Bezug auf Kinder, die wegen Fremdsprachigkeit über einen geringen Wortschatz verfügen.

Die KAB-C–Aufgaben sind für die Kinder sehr motivierend. Wie die Leistungen zeigten, ist die Eichung für eine Schweizer Stichprobe nicht ganz adäquat. Trotzdem können diese Aufgaben für die Messung von Mathematik Leistungen sehr empfohlen werden, die Aufgaben erwiesen sich als sehr altersgerecht.

Die Aufgabe zum Balancieren müsste überprüft werden. Wie wir gesehen haben sind die Ergebnisse leider nur teilweise interpretierbar. Hier zeigt sich, dass eine Voruntersuchung mit 32 Kindern keine ausreichende Sicherheit bietet, um die Aufgabenschwierigkeit zu eichen. Diese Aufgabe müsste nochmals überarbeitet werden, bevor sie weiter eingesetzt werden könnte.

Die Aufgabe zum Hüpfen hat sich gut bewährt, sie lässt sich auch auf kleinem Raum mit jüngeren Kindern ohne Gefahr durchführen.

Die Anwendung von adaptiven Aufgaben hat den Vorteil, dass die Schwierigkeiten in jedem Jahr vergleichbar sind. Allerdings machen Kinder ja nicht in jedem Jahr die gleichen Fortschritte. Beim KTK zeigte sich deutlich, dass die Normierung für eine Untersuchung über Pilotstudien abgesichert werden muss.

5. Literaturverzeichnis

- Achermann, E. (1992). *Mit Kindern Schule machen*. Zürich: Verband Lehrerinnen Zürich.
- Aiken, L. R. (1976). Update on attitudes and other affective variables in learning mathematics. *Review of Educational Research*, 46, 293-311.
- Althaus, L., & Bühler, G. (1999). *Aspekte der Motorik von Kindergartenkindern*. Unveröffentlichte Diplomarbeit. Bern: Institut für Sportwissenschaften der Universität Bern.
- Asendorpf, J.B., & van Aken, M.A.G. (1993). Deutsche Versionen der Selbstkonzeptskalen von Harter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25 (1), 64-86.
- Bandura, A. (1995). *Self-efficacy in Changig Societies*. New York: Cambridge University Press.
- Beneson, J.F., & Dweck, C.S. (1986). The development of trait explanations and self-evaluations in the academic and social domains. *Child Development*, 57, 1179-1187.
- Block, J., & Colvin, C.R. (1994). Positive Illusions and Well-being Revisited: Separating Fiction from fact. *Psychological Bulletin*, 116, 28-29.
- Buff, A. (1991). Schulische Selektion und Selbstkonzeptentwicklung. In R. Pekrun, & H. Fend, *Schule und Persönlichkeitsentwicklung* (S.100-125). Stuttgart: Enke.
- Catani, R. (1998a): *Kulturelle Psychologie und ihre Bedeutung für die Selbst- und Sozialkompetenz*. Unveröffentlicher Forschungsbericht. Spiez: Staatliches Seminar.
- Catani, R. (1998b): *Die untergeordnete Rolle der Frau in unserer Kultur*. Unveröffentlicher Forschungsbericht. Spiez: Staatliches Seminar.
- Catani, R. (1999): Didaktik im Wandel: Verpasste Entwicklungen und Möglichkeiten einer Neuausrichtung der Allgemeinen Didaktik. In M. Kübler, & S. Albisser (Hrsg), *Entwicklungen und Umbrüche in der Lehrerinnen- und Lehrerbildung* (S.103-118). Spiez: Staatliches Seminar.
- Chapman, J.W., & Tunmer, W.E. (1995). Development of Young Childrens Reading Self- Concepts: An Examination of Emerging Subcomonents and Their Relationships with Reading Achievement. *Journal of Educational Psychology*, 87, 154-167.
- Cole, D.A. et al. (1998). Children's Over- and Underestimation of Academic Competence: A Longitudinal Study of Gender Differences, Depression and Anxiety. *Child Development*, 70, 459-473.
- Das Zahlenbuch (1995). Stuttgart.Klett.
- Demény, A. (2002). *Selbsteinschätzung von Kindern und Jugendlichen im Bereich der*

- schulischen Leistung*. Lizentiatsarbeit. Universität Bern. Bern.
- Der Bund. *Neue Beurteilung für die Volksschulstufe*. Bern. Verlag der Bund. 153Jg. (Ausgabe 138), 2.
- Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.
- Eccles, J.S. (1984). Sex differences in mathematics participation. In M. Steincamp, & M.L. Maehr (Eds), *Advances in motivation and achievement (Vol.2)*, (pp97-132). Greenwich, Connecticut: JAI press.
- Eccles, J.S., Wigfield, A., Harold, R., Blumenfeld, P. (1993). Age and gender differences in childrens achievement self-perceptions during the elementary school years. *Child Development*, 64, 830-847.
- Erziehungsdirektion des Kantons Bern (1973): *Lehrplan für die Primarschule*. Bern: Staatlicher Lehrmittelverlag.
- Erziehungsdirektion des Kantons Bern (1982): *Lehrplan für die Primarschule*. Bern: Staatlicher Lehrmittelverlag.
- Erziehungsdirektion des Kantons Bern (1995): *Lehrplan für die Volksschule*. Bern: Staatlicher Lehrmittelverlag.
- Erziehungsdirektion des Kantons Bern (1997). Der Spezialunterricht im Kindergarten und in der Volksschule. In: *Richtlinien und Grundsätze der Erziehungsdirektion vom 24.3.1997*. Bern. Erziehungsdirektion.
- Erziehungsdirektion des Kantons Bern (1999). *Lehrplan Kindergarten*. Bern: Staatlicher Lehrmittelverlag.
- Erziehungsdirektion des Kantons (2003). *Fremdsprachige Kinder und Jugendliche in der Volksschule des Kantons Bern*. Bern.
- Eshel, Y., & Klein, Z. (1981): Development of academic self-concept of lower-class and middle-class primary school children. *Journal of Educational Psychology*, 73, 287-293.
- Faust-Siehl, G. et al. (1996). *Die Zukunft beginnt in der Grundschule*. Reinbeck bei Hamburg: Rowohlt.
- Fend, H. (1974). *Gesellschaftliche Bedingungen schulischer Sozialisation*. Weinheim und Basel: Beltz.
- Flammer, A. (1990). *Erfahrung der eigenen Wirksamkeit*. Bern: Huber.
- Flammer, A. (1995a). Possum, ergo sum – Nequeo sum qui sum. In F. Oosterwegel, Wicklund, R.A. (Eds), *The Self in European and North American Culture: Development and Processes* (pp333-349). Dordrecht: Kluwer.
- Flammer, A. (1995b). Developmental analysis of control beliefs. In A. Bandura (Ed), *Self-efficacy in Changing Societies* (pp69-113). New York: Cambridge

University Press

- Flammer, A. (1996). *Das kompetente Selbst und seine Entwicklung*. Vierteljahresschrift für Heilpädagogik und ihre Nachbarwissenschaften, 65, 266-279.
- Flammer, A. et al. (1995). Coping with control-failure in Japanese and Swiss adolescents. *Swiss Journal of Psychology*, 54, 277-288.
- Flammer, A., & Alsacker, F. D. (2002). *Entwicklungspsychologie der Adoleszenz*. Bern: Huber.
- Frey, K.S., & Ruble, D.N. (1987). What children say about classroom performance: sex and grade differences in perceived competence. *Child Development*, 58, 1066-1078.
- Gächter, E. (1988). Die Quartiere der Stadt Bern und ihre Struktur – eine Untersuchung mit quantitativen Methoden. *Berner Geografische Mitteilungen*, 16, 51-70.
- Gaschler, P. (1998). Motorische Entwicklung und Leistungsfähigkeit von Vorschulkindern- Empirische Untersuchung. *Haltung und Bewegung*, 4, 5-18.
- Gasser, P. (1995). *Neue Lernkultur*. Gerlafingen: Selbstverlag.
- Geppert, U. (1997). Entwicklung lern- und leistungsbezogener Motive und -einstellungen. In F.E. Weinert, & A. Helmke (Hrsg). *Entwicklung im Grundschulalter* (S.43-82). Weinheim: Psychologie Verlagsunion.
- Geppert, U., & Heckhausen, H. (1990). Ontogenese der Emotion. In K.R. Scherer (Hrsg). *Psychologie der Emotion. Enzyklopädie der Psychologie. Bd. C/IV/3*, 115-213. Göttingen: Hogrefe.
- Giesecke, H. (1988). *Das Ende der Erziehung*. Stuttgart: Klett.
- Grob, A. (1996). Entwicklung und Regulation des subjektiven Wohlbefindens. Integrierendes Manuskript zur Erlangung der Habilitation. Universität Bern. Bern.
- Hansford, B.C., & Hattie, J.A. (1982). The relationship between self and achievement / performance measures. *Review of Educational Research*, 52, 123-142.
- Harter, S. (1982). The perceived competence scale for children. *Child Development*, 53, 87-97.
- Harter, S. (1983). Developmental perspectives on the self-esteem. In: P.H. Mussen, *Handbook of child psychology. Vol.4*, 275-385. New York: Wiley.
- Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child development*, 55, 1969-1982.
- Heckhausen, H. (1974). *Leistung und Chancengleichheit*. Göttingen: Hogrefe.
- Heckhausen, H. (1980). *Motivation und Handeln*. Heidelberg: Springer.
- Heckhausen, H., & Rheinberg, F. (1980). Lernmotivation im Unterricht, erneut betrachtet. *Unterrichtswissenschaft*, 8, 7-47.
- Helmke, A. (1991). Entwicklung des Fähigkeitsselbstbildes vom Kindergarten bis zur dritten Klasse. In R. Pekrun, & H. Fend, *Schule und Persönlichkeitsentwicklung* (S.83-99).

- Stuttgart: Enke.
- Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen*. Göttingen: Hogrefe.
- Helmke, A. (1993). Die Entwicklung der Lernfreude vom Kindergarten bis zur 5.Klassenstufe. *Zeitschrift für Pädagogische Psychologie*, 7, 77-86.
- Helmke, A., & Van Aken, M. (1995): The Causal ordering of academic achievement and self-concept ability during elementary school: a longitudinal study. *Journal of Educational Psychology*, 87, 624-637.
- Helmke, A., & Weinert, A. (1997). *Entwicklung im Grundschulalter*. Weinheim: Psychologie Verlagsunion.
- Ingold, A. (2002). *Selbstüberschätzung der eigenen physischen Leistungsfähigkeit*. Lizentiatsarbeit. Universität Bern. Bern.
- Jacazinski, C.M., & Nicholls, J.G. (1990). Reducing effort to protect perceived ability. *Journal of Educational Psychology*, 82,15-21.
- Jerusalem, M., & Schwarzer, R. (1991). Entwicklung des Selbstkonzepts in verschiedenen Lernumwelten. In: R. Pekrun, H. Fend, (Hrsg). *Schule und Persönlichkeitsentwicklung* (S.115-130). Stuttgart: Enke.
- Kanton Bern (1992). *Gesetz über die Volksschulen des Kantons Bern*. Bern: Staatskanzlei.
- Kelley, H.H. (1973). The process of causal attribution. *American Psychologist*. 28, 107-128.
- Kiphard, E.J., & Schilling, F. (1974). *Körperkoordinationstest für Kinder*. Weinheim. Beltz.
- Kramer, J. (1972). *Kramer-Test*. Solothurn: Antonius-Verlag.
- Kun, A. (1977). Development of the Magnitude-covariation and compensation schemata in ability and effort attributions of performance. *Child Development*, 48, 863-873.
- Little, T., Oettingen, G., Stetsenko, A., & Baltes, P.B. (1995). Childrens action-control-beliefs and school performance: How do American children compare with German and Russian children? *Journal of Personality and Social Psychology*, 69, 686-700.
- Lobeck, A. (1987). *Rechentest 2.Klasse*. Basel: Beltz.
- Markus, H.R., Mullally, P.R., Kitayama, S. (1997). Selfways in modes of cultural participation. In U. Neisser, D. Joplins, (Eds), *The conceptual self in context: Culture, experience, self-understanding* (pp13-61). New York: Cambridge University Press.
- Markus, H.R., Kitayama, S., Matsumoto, H. & Norasakuhnit, U. (1997). Individual and collective processes in the consturction of the self: Self – enhancement in the United States and self –criticism in Japan. *Journal of Personality and Social Psychology*, 58.
- Measelle, J.R. et al. (1998). Assessing Young Children's Views of Their Academic, Social, and Emotional Lives: An Evaluation of the Self-Perception Scales of the Berkeley Puppet Interview. *Child Development*, 69, 1556-1576.
- Meier, A. (1999). Psychologie in der Primarlehrerinnen- und Lehrerausbildung. In M.Kübler, S. Albisser (Hrsg), *Entwicklungen und Umbrüche in der Lehrerinnen- und*

- Lehrerbildung* (S.119-138). Spiez: Staatliches Seminar.
- Melchers, P., & Preuss, U. (1991). *Kaufman-Assessment-Battery for Children*. Niederlande: Swets & Zeitlinger.
- Meyer, W.U. (1984). Das Konzept von der eigenen Begabung. *Psychologische Rundschau*, 35, 136-150.
- Mietzel, G. (1993). *Psychologie in Unterricht und Erziehung*. Göttingen: Hogrefe.
- Marsh, H.W., Barnes, J., Cairns, L., Tidman, M. (1984). Self-Description Questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology*, 76, 940-956.
- Marsh, H.W., Craven, R., Debus, R. (1998). Structure, Stability, and Development of Young Children's Self-Concepts: A Multicohort-Multioccasion Study. *Child Development*, 69, 1030-1053.
- Nicholls, J.G. (1978). The Development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development*, 49, 800-814.
- Nicholls, J.G. (1979). Development of perception of own attainment and causal attributions for success and failure in reading. *Journal of Educational Psychology*, 71, 94-99.
- Nicholls, J.G., Miller, A. (1983). The differentiation of the concepts of difficulty and ability. *Child Development*, 54, 951-959.
- Nicholls, J.G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91, 328-346.
- Nicholls, J.G., Miller, A.T. (1984). Development and its discontents. In *Advances in motivation and achievement*, 3, 185-218. JAI press.
- Oettingen, G., Lindenberger, U., Baltes, P.B. (1992). Sind die schulleistungsbezogenen Überzeugungen der Ostberliner Kinder entwicklungshemmend? *Zeitschrift für Pädagogik*, 2, 299-324.
- Oerter, R., & Montada, L. (1995). *Entwicklungspsychologie*. Weinheim: PVU.
- Pekrun, R., & Fend, H. (1991). *Schule und Persönlichkeitsentwicklung*. Stuttgart: Enke.
- Petillon, H. (1993). *Das Sozialleben des Schulanfängers*. Weinheim: PVU.
- Pfister, S. (1992). Kontrollmeinungen und Kontrollmeinungsbegründungen von acht- und neunjährigen Kindern. Lizentiatsarbeit. Bern. Universität Bern.
- Reble, A. (1981). *Geschichte der Pädagogik*. Frankfurt am Main. Klett-Cotta.
- Rheinberg, F., & Einstrup, B. (1977). Selbstkonzept der Begabung bei Normal- und Sonderschülern gleicher Intelligenz: ein Bezugsgruppeneffekt. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 9, 171-180.
- Ruble, D.N., Boggiano, A.K., Feldman, N.S., Loebel, J.H. (1980). Social comparison in Self-evaluation. *American Psychologist*, 105-116.

- Ruble, D.N., & Flett, G.L. (1988). Conflicting goals in self-evaluative information seeking: Developmental and ability level analyses. *Child Development*, 59, 97-106.
- Ruble, D.N., Parsons, J.E., & Ross, J. (1976). Self-evaluative responses of children in an achievement setting. *Child Development*, 48, 990-997.
- Ruf, U., & Gallin, P. (1990). *Sprache und Mathematik in der Schule*. Zürich: Verband Lehrerinnen Zürich.
- Rustmeyer, R., Jubel, A. (1996). Geschlechtsspezifische Unterschiede im Unterrichtsfach Mathematik hinsichtlich der Fähigkeitseinschätzung, Leistungserwartung, Attribution sowie im Lernaufwand und im Interesse. *Zeitschrift für Pädagogische Psychologie*, 10,(1), 13-25.
- Schweizerische Koordinationsstelle für Bildungsforschung (SKF), (1999). *Begabungsförderung in der Volksschule-Umgang mit Heterogenität. Trendbericht*. Aarau: SKF.
- Seminar Spiez (2000). Beobachtungsberichte zur Psychologie des Lernalters. LAB 9. Spiez. Unveröffentlicht. Staatliches Seminar.
- Shaklee, H. (1976). Development in Inferences of Ability and Task Difficulty. *Child Development*, 47, 1051-1057.
- Shultz, T.R., Butkowsky, I., Pearce, J.W., Shanfield, H. (1975). Development of schemes for the attribution of multiple psychological causes. *Developmental Psychology*, 11, 502-510.
- Staatlicher Lehrmittelverlag (2000). *Sprachfenster*. Bern.
- Stipek, D.J., & Daniels, D.H. (1988). Declining perceptions of competence. *Journal of Educational Psychology*, 80, 352-356.
- Stipek, D.J., & DeCotis, K.M. (1988). Children's Understanding of Implications of Causal Attributions for Emotional Experience. *Child Development*, 59, 1601-1610.
- Stipek, D.J., & MacIver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development*, 60, 521-538.
- Stipek, D.J., & Tannatt, L.M. (1984). Children's judgements of their own and their peers' academic competence. *Journal of Educational Psychology*, 76, 75-84.
- Taylor, S.E., & Brown, J.D. (1994). Positive Illusions and well-being Revisited: Separation fact from fiction. *Psychological Review*, 116, 21-27.
- Tiedemann, J. (1995). Geschlechtstypische Erwartungen von Lehrkräften im Mathematikunterricht der Grundschule. *Zeitschrift für Pädagogische Psychologie*, 9, 153-161.
- Tiedemann, J., & Faber, G. (1994). Mädchen und Grundschulmathematik: Ergebnisse einer vierjährigen Längsschnittuntersuchung zu ausgewählten geschlechtsbezogenen Unterschieden in der Leistungsentwicklung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 26, 101-111.
- Tiedemann, J., & Faber, G. (1995). Mädchen im Mathematikunterricht: Selbstkonzept und

- Kausalattributionen im Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27, 61-71.
- Thomas, J.R., & French, K. E. (1985). Gender Differences across age in major motor Performance: A meta-analysis. *Psychological Bulletin*, 98, 260-282.
- Valeski, T.N., & Stipek, D.J. (2001). Young childrens feelings about school. *Child Development*, 72, 1198-1213.
- Valkanover, S. (2003). *Aspekte der Körpererfahrung und der Motorik im Kontext von Mobbing im Kindergarten*. Dissertation. Universität Bern. Bern.
- Weidenmann, B., & Krapp, A. (1986). *Pädagogische Psychologie*. Weinheim: PVU.
- Weinert, F.E., & Stefanek, J. (1997). Entwicklung vor und während der Grundschulzeit. In F.E. Weinert, & A. Helmke, *Entwicklung im Grundschulalter* (S.423-452). Weinheim: Psychologie Verlags Union.
- Wegge, J. (1998). Lernmotivation, Informationsverarbeitung, Leistung. In D.H. Rost, *Pädagogische Psychologie und Entwicklungspsychologie, Bd.7*. Münster: Waxmann.
- Wicki, W., Reber, R., Flammer, A., Grob, A. (1994). Begründung der Kontrollmeinung bei Jugendlichen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 26, 243-261.
- Wigfield, A., Eccles, J.S., Mac Iver, D., Reuman, D., Midgley, C. (1991). Transitions at early adolescence: Changes in childrens domainspecific self-perceptions and general self-esteem across the transition to junior high school. *Development Psychology*, 27, 552-565.
- Wigfield, A., Eccles, J.S., Harold, R.D., Freedman-Doan, C., Blumenfeld, P.C. (1997). Change in childrens competence belief and subjective task values across the elementary school years: a 3-year study. *Journal of Educational Psychology*. 89, 451-469.

6. Anhang

Instruktionsmanual für Untersuchungsdurchführende

Informationen an Behörden, Lehrpersonen

Einschätztreppe in Originalgrösse, einfarbig (im Experiment farbig)

Befragung und Untersuchung Datum: _____ Uhrzeit: _____
Befragung von Kindern zur Leistungsfähigkeit im Kindergartenalter 31.5.99 /R.Catani

Ort: _____ Kindergärtnerin: _____ Untersucht von: _____

1. Einleitung

Ich heiße und wie heißt du? wir machen jetzt zusammen ein paar Aufgaben, versuche diese Aufgaben so gut zu machen, wie du kannst. Immer bevor du die Aufgaben machst, musst du schätzen, wie gut du solche Aufgaben kannst.

Damit ich nicht vergesse, was du gesagt hast, schreibe ich es mir auf,

hier deinen Vor - und Familiennamen: _____

oder ich mache ein Kreuz und zwar so : weiblich: _____ Männlich: _____

(Geburtsdatum / Geburtstag bei Kindergärtnerin erfragen) _____

In einer Aufgabe geht es darum , einen Schaumgummiblock zu überspringen, bei einer anderen Aufgabe muss man über ein schmales Brett gehen !

Bei Kindern gibt es ja solche, die besser über ein schmales Brett gehen können oder über ein Hindernis springen können und solche, die das nicht so gut können.

Treppenkarte (die Kinder auf den Stufen jeweils zeigen mit dem Finger)

Genau das siehst du hier auf dem Bild: hier in der Mitte sind die Kindergärteler, sie haben die Farbe Viele können *manchmal gut und manchmal nicht so gut* springen oder über einen Balken gehen , andere sind *recht gut, gut* oder *sehr gut* und sogar *besser als* die Erstklässler, die hier oben stehen!

Je besser, desto weiter oben, ja es geht natürlich noch weiter als bis zu den Erstklässlern! Andererseits gibt es solche, die können es *noch nicht so gut* oder *gar nicht gut*, ja die können es sogar schlechter als Kinder, die noch nicht im Kindergarten sind und hier unten sind, aber nächstes Jahr sind sie dann vielleicht auch hier !

Also hier in der Mitte sind die Kindergärteler die das Ueberspringen von Schaumgummiblocks *manchmal gut und manchmal nicht so gut* können und wenn du jetzt so schaust, wo wärst denn du auf dieser Treppe ? (Buchstabe notieren) _____

(Antwort verbal verifizieren, beispielsweise: du meinst also, dass du bei den Kindern bist die das recht gut können ? / sehr gut können? / gar nicht gut können?

(Falls das Kind jetzt korrigiert, die neue Antwort wieder verbal bestätigen).

Stell Dir vor, dies wären die Kindergärteler, die das Rückwärtsgehen auf dem Brett ausprobiert haben. Einige konnten es *recht gut*, andere konnten es noch *nicht so gut*, oder sogar *schlechter als* die Kinder, die jetzt dann erst in den Kindergarten kommen. Hier sind dann die Kindergärteler, die das Rückwärtsgehen *sehr gut* konnten oder ein Kind war sogar *besser als* die Erstklässler.

In der Mitte sind wieder diejenigen, die es *manchmal gut* und *manchmal nicht so gut* können, sage mir jetzt, wie gut du rückwärts gehen kannst, was vermutes du, wo du stehen würdest auf dieser Treppe? (Buchstabe notieren) _____

2. Rückwärts balancieren: (man kann auch mit einer anderen Aufgabe beginnen)

Siehst du diese Balken? auf einer Seite sind sie breit, auf der anderen Seite schmal. Wenn man jetzt über diesen Balken läuft, dann wird es immer schwieriger, den Boden nicht mit dem Fuss zu berühren, weil er immer schmaler wird!

Du darfst nachher zweimal probieren über diesen Balken zu gehen, was glaubst du, wie weit du kommst ohne den Boden zu berühren, versuche so genau zu schätzen wie möglich !

Zeige mit dem Finger, bis zu welchem Strich du kommst ! (notieren) cm _____

Jetzt darfst du zweimal probieren !

erzieltes Resultat : 1.Versuch cm: _____ 2. Versuch : cm: _____

3. Ueberhüpfen auf einem Bein

jetzt sollst du dann über solche Schaumgummis hüpfen. Ich mache es vor (20 cm).
Versuche es gleich mal mit dem rechten (zeigen!) Bein über 1 Schaumgummi (Anlauf ca 1,5 Meter) und hüpf noch einen Schritt weiter. Falls du die Schaumgummis berührst, ist es ein Fehler, ebenso, wenn du mit beiden Beinen absteht,so (vorzeigen).

Versuchssprung des Kindes einmal mit jedem Bein.

"Schaumgummitreppe" zeigen: (zwei / zwei / eine / vier)

Instruktion: Die Kinder, die dann nächstes Jahr in den Kindergarten kommen, *springen meistens etwa* so hoch (2 zeigen), Kindergärteler, so wie du, *springen meistens etwa* so hoch (4 zeigen), 1.Klässler *springen meistens etwa* so hoch (5 zeigen), aber es gibt *ganz selten auch einen Knaben oder ein Mädchen* im Kindergarten, das so hoch springt (9 zeigen).

Was glaubst du, wie hoch kannst du ohne Fehler überspringen ? Anzahl Gummis _____

Welches Bein willst du zuerst versuchen ? _____ (links oder rechts?)

1. Versuch auf Höhe mit 1 Platte

2.Versuch: nur wenn 1. Versuch Fehler (touchieren/verschieben, beidbeinig abstehen)

3.Versuch: wenn dreimal nicht übersprungen, respektive Fehler, wird abgebrochen.

Wenn im 3. Versuch übersprungen wird, geht es weiter auf neuer Höhe

Anzahl Platten gültig mit einem Bein übersprungen: _____

und zwar im _____ Versuch.

Jetzt kommt das andere Bein dran !

Beginn: 1 Platte

Anzahl Platten gültig mit einem Bein übersprungen: _____
und zwar im _____ Versuch.

So, merci für deinen Einsatz !

4. Rechnen (-> Wenn in einer Aufgabengruppe keine Aufgabe richtig, zur Häslkarte gehen)

Treppenkarte

Bevor wir jetzt weiterfahren, zeige ich Dir aber wieder die "Stäge" /Treppe auf der die Kinder stehen. Wir sehen hier Kinder, sie immer besser werden, dieses Kind ist vielleicht in einem Jahr hier oder hier , es geht immer weiter.

Hier in der Mitte sind die Kindergärteler. Viele Kindergärteler sind *manchmal gut und manchmal nicht so gut* im Rechnen, andere sind *recht gut, gut* oder *sehr gut* und sogar *besser als* die Erstklässler!

Je besser, desto weiter oben, ja es geht natürlich noch weiter als bis zu den Erstklässlern! Andererseits gibt es Kinder, die können *noch nicht so gut* oder *gar nicht gut* rechnen, ja es gibt solche, die können sogar *schlechter rechnen* als Kinder, die noch nicht im Kindergarten sind! Die sind vielleicht erst in einem Jahr hier!

Also hier sind die Kindergärteler und wenn du jetzt so schaust, wo wärst denn du auf dieser Treppe ?

(Buchstabe notieren) _____ Interpretieren: du bist also bei den sehr guten/ ...

Karte 0

Stelle dir vor eine Familie geht in einen Zoo Tierpark: warst du auch schon einmal in einem solchen Park? Siehst du, das sind die Kinder der Familie!

Bei diesen Aufgaben geht es um das Rechnen, du musst Zahlen kennen und das Zählen: Wieviele Kinder hat es beispielsweise auf dieser Karte ? (Falls falsch, zählen lassen bis 3). Wir schauen uns jetzt diesen Zoo/Tierpark an.

(die Antwort aufschreiben, wenn keine Antwort: Querstrich)

Karte 1:

Aufgabe 6: Zeige auf die drei und sage: "was ist das für eine Zahl?" (drei)_____

Aufgabe 7: Zeige auf die sieben und sage: "und wie heisst diese Zahl?" (sieben)_____

Karte 2:

Aufgabe 8: Sage: "Zähle alle Robben" (sieben)_____
 (wenn das Kind "seehunde" oder so sagt, nicht korrigieren)

Karte 3:

Aufgabe 9: "Diese Leute warten in der Reihe, um Eintrittskarten zu kaufen.
 Zähle alle Leute in der Reihe" (neun)_____

Aufgabe 10a: "Sind mehr Leute in der Reihe oder hat es mehr Robben ?"(Mehr Leute)

Karte 4:

Aufgabe 10b: "Sind mehr Leute in der Familie oder hat es mehr Robben?"(mehr Robben)
 wenn 10 a und b richtig: Punkt_____

_____ weiterfahren, wenn mind. 1 Punkt, sonst Häslkarte ! _____

Karte 5:

Aufgabe 11: "Neun Kinder sind in der Reihe. Wenn fünf Kinder weggehen, wieviele Kinder bleiben dann übrig?" (vier) _____

Aufgabe 12: Zeige auf das erste Kind in der Reihe und sage:
"Dieses Kind ist das erste in der Reihe. Zeige mir bitte das dritte Kind in der Reihe." _____

Karte 6:

Aufgabe 13: Zeige auf das leere Kästchen: "Eine Zahl fehlt hier. Welche Zahl müsste in diesem Kästchen stehen?" (vier) _____

Karte 7:

Aufgabe 14a: "Schau dir die Leute an vor dem Affenkäfig. Sind da mehr Menschen oder mehr Affen oder gleichviele Menschen und Affen?" (gleichviel)

Karte 8:

Aufgabe 14b: "Sind auf diesem Bild mehr Menschen, mehr Affen oder gleich viel Menschen wie Affen auf dem Bild?" (gleichviel)
wenn a und b richtig: Punkt _____

_____weiterfahren,wenn mindestens 1 Punkt, sonst Hässlikarte ! _____

Karte 9:

Aufgabe 15: Auf diesem Bild steht, wieviele Tiere in diesem Monat im Zoo geboren wurden (Zeige auf die 10). Welche Zahl ist das ? (Antwortet das Kind "eins-null": "sag die Zahl anders") _____

Aufgabe 16: Zeige auf die 12 : " Welche Zahl ist das?"
(Antwortet das Kind "1-2",dann: sag die Zahl anders"). _____

Aufgabe 17: Zeige auf die 17: Welche Zahl ist das ?"
(antwortet das Kind...siehe oben) _____

Karte 10:

Aufgabe 18: "Eine Familie geht zu den Elefanten. Hier siehst du 6 Elefanten. Wenn vier dieser Elefanten weggehen würden, wieviele bleiben dann noch übrig?" (zwei) _____

_____weiterfahren,wenn mindestens 1 Punkt, sonst Hässlikarte ! _____

Karte 11:

Aufgabe 19: "Auf diesem Schild steht das Alter der Elefanten. (Zeige auf 37). Welche Zahl ist das ?
(Antwortet das Kind "Drei -sieben", dann....) _____

Aufgabe 20: Schau dir diese Zahl (29) genau an: Wenn man zählt, welche Zahl kommt nach dieser Zahl ?" (dreissig)

Karte 12:

Aufgabe 21: "Wenn fünf Elefanten zu den sechs hier auf dem Bild hinzukommen, wieviele sind es dann alle zusammen?" (elf)

_____ weiterfahren, wenn mindestens 1 Punkt, sonst Hässlikarte ! _____

Karte 13:

Aufgabe 22: "Dieser Mann verkauft Erdnüsse, damit die Besucher des Zoos mit den Nüssen die Elefanten füttern können. Hier ist ein Schild (zeigen). Dreissig Rappen kostet ein Beutel Nüsse, wieviel kosten zwei Beutel ?" (60) _____

Karte 14:

Aufgabe 23: "Hier sind wir beim Kinderzoo angekommen. Zeige mir die Hälfte der Schäfchen. Zeige und berühre sie mit deinem Finger.(4) _____
(Das Kind darf die Tiere einzeln zählen oder einkreisen oder abdecken)

Karte 15:

Aufgabe 24: "Der Zoo hat doppelt so viele Giraffen wie Geissen. Es gibt im Zoo fünf Geissen. Wieviele Giraffen hat es dann ?" (zehn) _____

Karte 16:

Aufgabe 25: "Die Kinder gehen zu den Kaninchen, um sie zu streicheln. Sie teilen die Kaninchen gleichmässig untereinander auf. Wieviel Kaninchen bekommt jedes Kind zum Streicheln?" (vier) _____

_____ weiterfahren, wenn mindestens 1 Punkt, sonst Hässlikarte ! _____

Karte 17:

Aufgabe 26: "Mit welcher Zahl muss man die Anzahl der Enten malnehmen, um genau soviel Enten wie Schäfchen zu haben ?" (zwei) _____

Karte 18:

Aufgabe 27: "Jetzt geht die Familie nochmals zu den Elefanten. Dieser grosse Elefant (zeigen) ist 650 Kilogramm schwer. Der kleine hier (zeigen) ist 550 Kilogramm schwer. Wieviel ist der grössere Elefant schwerer, wieviele Kilogramm wiegt er mehr?" (hundert) _____

Aufgabe 28: "Wieviel wiegen der grosse und der kleine Elefant zusammen? -> Erwinnere dich, der grosse wiegt 650 Kilogramm und der kleine wiegt 550 Kilogramm." (1200Kg 1t200kg) _____

_____ weiterfahren, wenn mindestens 1 Punkt, sonst Hässlikarte ! _____

Karte 19:

Aufgabe 29: "Die Familie will nochmals Nüsse kaufen, um Elefanten zu füttern. Ein Beutel mit Nüssen kostet dreissig Rappen. wieviel kosten 7 Beutel? (2.10) _____

Karte 20:

Aufgabe 30: "Dieses Schild sagt etwas über die Fütterung der Robben. Der Eintritt für einen Erwachsenen kostet 75 Rappen. Der eintritt für Kinder kostet ein Drittel des Eintritts für Erwachsene.

Wieviel kostet der Eintritt für Kinder ?" (25 Rappen) _____

Karte 21:

Aufgabe 31: "80 Personen gehen zur Robbenfütterung. Alle passen in vier Sitzreihen hinein. Wenn in jeder Sitzreihe die gleiche Anzahl Personen sitzt, wieviel Personen sitzen dann jeweils in einer Sitzreihe ?" (zwanzig) _____

Karte 22:

Aufgabe 32: "Hier steht ein anderes Schild für die Robbenfütterung. Jede Robbenfütterung beginnt genau zu der auf der Tafel angegebenen Zeit und dauert genau 75 Minuten. Wieviel Zeit ist jeweils zwischen dem Ende einer und dem Beginn der nächsten Fütterung?" (15 min.) _____

Häsilkarte Karte 16) :

Zähle alle weissen Häschen ! _____

Wieviele Schnautzhaare hat das braunweisse Häschen (zeigen) _____

Wieviele Ohren haben Hasen ? und wieviele hast du? wer hat jetzt mehr ? _____

5. Abschluss

So , jetzt hast du dir aber schön Mühe gegeben bei diesen Aufgaben !

Merci fürs Mitmachen ! Du kannst jetzt wieder zu den anderen, aber verrate nicht was wir gemacht haben, sonst ist es für die anderen nicht so spannend!

(zurückbringen und das nächste Kind holen !)

Reto Catani
Seminarlehrer
Hohlenhausweg 9
3672 Oberdiessbach
Tel.: 031 771 15 87

Oberdiessbach, im Juni 1999

Sehr geehrte Damen und Herren

Neben meiner Tätigkeit als Ausbildner von Lehrkräften am Seminar Spiez, mache ich zur Zeit eine wissenschaftliche Untersuchung zur Entwicklung der Fähigkeit von Kindern, sich in bestimmten Leistungsbereichen selber einschätzen zu können.

Die Kindergärtnerin Ihres Kindes hat sich bereit erklärt, mit Ihrer Klasse bei dieser Untersuchung mitzuwirken.

Für Ihr Kinder bedeutet das, dass an einem von der Kindergärtnerin angekündigten Tag, Lehrerstudenten im Kindergarten diese Untersuchung durchführen werden.

Eine Studentin wird dabei Ihr Kind zu drei Aufgaben befragen. Das Kind soll vermuten, wie weit es über ein Brett balancieren kann, ohne das Gleichgewicht zu verlieren.

In einer zweiten Aufgabe geht es darum, über einen Schaumgummiblock zu springen und in der dritten Aufgabe geht es um einfache Zählaufgaben.

Sobald das Kind die Aufgaben bewältigt hat, wird es von der Studentin wieder in die Klasse zurückgebracht. Alle Aufgaben sind speziell für diese Altersgruppe konstruiert worden.

Die Untersuchung wird in einem Jahr und in zwei Jahren wiederholt. Ich möchte so herausfinden, ob sich Kinder vom Kindergarten bis zur zweiten Klasse immer genauer einschätzen können. Wie Sie vielleicht wissen, müssen sich die Kinder ab der dritten Klasse im Lernbericht selber einschätzen können. Meine Untersuchung soll dazu beitragen, herauszufinden, wie gut das den Kindern in der Unterstufe gelingt. Leider gibt es dazu weder in der Schweiz noch sonst in Europa oder Amerika Forschungsergebnisse. Hier schliesst meine Forschung eine wichtige Lücke.

Ich wäre deshalb sehr froh, wenn Ihr Kind auch mitmacht, damit wir möglichst über 200 Kinder einbeziehen können.

Da wir den Namen Ihres Kindes aus Datenschutzgründen verschlüsseln, können wir Ihnen dann leider nicht sagen, wie Ihr Kind sich im Vergleich zu den anderen eingeschätzt hat.

Wir werden Sie aber nach Abschluss der Untersuchung, also wenn Ihr Kind in die dritte Klasse geht, über die Ergebnisse der Gesamtuntersuchung orientieren.

Die Erziehungsdirektion und die Inspektorate unterstützen solche Untersuchungen, weil sie beispielsweise für die Erarbeitung von neuen Lehrmitteln oder Lehrplänen wichtig sind.

Die zuständige Kindergarten- oder Schulkommission ist über diese Untersuchung informiert.

Selbstverständlich steht es Ihnen zu, Ihr Kind von der Untersuchung zu dispensieren.
Falls Sie das wünschen, dann lassen Sie diesen Brief durch Ihr Kind unterschrieben der
Kindergärtnerin zukommen.

Ich wäre Ihnen sehr dankbar, wenn Ihr Kind auch mitmachen könnte und verbleibe:

Mit freundlichem Gruss:

Reto Catani

Ich wünsche, dass mein Kind nicht mit macht: Unterschrift: _____
Reto Catani

Seminarlehrer Spiez
Hohlenhausweg 9
3672 Oberdiessbach
Tel.: 031 / 771 15 87

8. Juni 1999

Liebe Kolleginnen

Nun ist es also so weit, wir kommen am folgenden Tag, um die Untersuchung zu machen:

Datum : _____ Zeit: _____

Je nach Anzahl Kindern kommen wir also etwas vor dem offiziellen Kindergartenbeginn, sodass wir sicher alle Kinder untersuchen können. Pro Kind dauert die Untersuchung knapp 30 Minuten.

Wichtig ist, dass Sie sich überlegen, wo wir die Untersuchung genau durchführen können. Es braucht nicht viel Platz, aber wichtig ist, dass die Klassenarbeit und die Untersuchung sich nicht gegenseitig stören. Sofern es nicht regnet, kann die Untersuchung auch im Garten stattfinden. Folgender Platzbedarf ist einzurechnen:

- a) Rückwärts balancieren: vier mal ein Meter
- b) überspringen des Schaumgummis: drei mal ein Meter
- c) Rechnen: Ecke mit zwei Sitzgelegenheiten

Für die Untersuchung brauchen wir die Geburtsdaten der Kinder. Erfahrungsgemäss fragen wir dazu die Kindergärtnerinnen, wir sind also froh, wenn Sie die Daten griffbereit haben, damit wir die Geburtsdaten am Mittag eintragen können.

Die Untersuchung wird von einem Team durchgeführt, welches aus Studierenden meines Ausbildungskurses besteht. Die Studierenden empfangen in der gleichen Woche ihr Lehrpatent, sie sind also den Umgang mit Kindern gewohnt und sollten deshalb keine Schwierigkeiten bei der Betreuung der Kinder haben.

Ich möchte Ihnen hiermit nochmals herzlich danken, dass diese Untersuchung gemacht werden darf und danke Ihnen im voraus für ihre Bemühungen!

Falls die Studierenden noch Fragen bezüglich dem Ort ihres Kindergartens haben, werden sie telefonisch mit Ihnen Kontakt aufnehmen.

Mit freundlichem Gruss

Reto Catani

