

The final publication is available at Hogrefe via DOI:

<http://dx.doi.org/10.1026/0049-8637.36.1.1>

Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 36, 1-9 © 2004 by Dickhäuser, O. & Galfe, E.

This version of the article may not completely replicate the final version published in Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie. It is not the version of record and is therefore not suitable for citation.

Kolummentitel: VERGLEICHSPROZESSE UND FÄHIGKEITSELBSTKONZEPT

Besser als..., schlechter als...

Leistungsbezogene Vergleichsprozesse in der Grundschule

Better..., worse...

Achievement related comparison processes in Elementary School

Oliver Dickhäuser & Eva Galfe

Justus-Liebig-Universität Gießen

Summary

Social, dimensional and temporal comparison processes are mediators between achievement and the self-concept of ability. This study reconsiders the relationship between comparison processes and the self-concept of ability (SCA) in elementary school. It also considers performances, which contradict students' achievement expectations. The day before a math-exam 379 third- and fourth-graders were asked to report their expectations with regard to this exam and their mathematical SCA. Immediately following the announcement of their exam results the students answered questions concerning the comparison-processes and again concerning their SCA. As expected comparisons with better-performing students (upward social comparisons) have negative effects on the SCA, whereas comparisons with students, who have done worse in the exam (downward social comparisons), have positive effects. In spite of the negative consequences of social upward-comparisons students are more likely to engage in this kind of comparison than in downward-comparisons, especially if their performance is worse than they have expected. Dimensional and temporal comparisons did not affect the SCA.

Key words: Self-concept - ability - mathematical achievement - comparison-processes - expectation - elementary school

Zusammenfassung

Soziale, dimensionale und temporale Vergleichsprozesse stellen vermittelnde Mechanismen zwischen erbrachter Leistung und dem Fähigkeitsselbstkonzept (FSK) dar. Die vorliegende Arbeit überprüft den Zusammenhang zwischen Vergleichsprozessen und dem mathematischen FSK im Grundschulalter. Dabei wird berücksichtigt, ob die Leistung den Erwartungen der Schüler/-innen widerspricht. 379 Dritt- und Viertklässler wurden am Tag vor einer Mathematikarbeit zu ihrer Notenerwartung und zu ihrem mathematischen FSK befragt. Nach der Rückgabe der benoteten Arbeit machten die Schüler Angaben zu den Vergleichsprozessen sowie erneut zum FSK. Wie erwartet zeigte sich, dass Vergleiche mit besseren Mitschülern (soziale Aufwärtsvergleiche) negative Auswirkungen auf das FSK haben, soziale Abwärtsvergleiche dagegen positive. Trotz der negativen Konsequenzen sozialer Aufwärtsvergleiche werden diese gegenüber Abwärtsvergleichen bevorzugt, insbesondere, wenn die Leistung schlechter als erwartet ausfiel. Dimensionale und temporale Vergleiche haben den Befunden zufolge keine Effekte auf das FSK.

Schlagwörter: Selbstkonzept - Fähigkeit - Mathematikleistung - Vergleichsprozesse - Erwartung - Grundschule

Vorstellungen, die Personen über die Höhe eigener Fähigkeiten haben (Fähigkeitsselbstkonzepte), sind ein bedeutsamer Untersuchungsgegenstand der Psychologie. Dies hat verschiedene Gründe. Zum einen ist die Höhe der wahrgenommenen eigenen Fähigkeiten für Individuen subjektiv von hoher Bedeutung, da eigene Kompetenz vielfach als positiv, eigene Unfähigkeit als negativ gewertet wird (vgl. Meyer, 1984). Darüber hinaus konnte gezeigt werden, dass Fähigkeitsselbstkonzepte (FSK) sich in vielfältiger Weise auf das Erleben und Verhalten von Personen in Leistungssituationen auswirken können. So beeinflusst das FSK unter bestimmten Bedingungen Erfolgserwartungen in Leistungssituationen, Attributionen, Wahlverhalten und (nachfolgend) Leistung (zsf. Wigfield & Karpathian, 1991).

Bei Schülern sind die Vorstellungen über die Höhe eigener Fähigkeiten fachspezifisch differenziert. So können klar ein mathematisches und ein verbales FSK unterschieden werden (Marsh & Shavelson, 1985). Eine hinter dieser Unterscheidung liegende Annahme ist, dass das (fach-)spezifische FSK häufig in engerer Beziehung zum konkreten Verhalten einer Person steht als das globale Selbstkonzept. Will man die Entstehung von Fähigkeitsselbstkonzepten erklären, ist es sinnvoll, zunächst Einflüsse auf Ebene des fachspezifischen oder aufgabenspezifischen FSK zu untersuchen (siehe z.B. Möller, 2000).

Der skill-development-Ansatz nimmt an, dass die Leistung, die eine Person erbringt, eine wesentliche Determinante des FSK darstellt (vgl. Helmke, 1992). In pfadanalytische Studien

(z.B. Helmke & Aken, 1995) zeigt sich dabei, dass Noten ein besserer Prädiktor von Fähigkeitsselbstkonzepten sind als standardisierte Testscores (vgl. auch Aken, Helmke & Schneider, 1997). Dies liegt vermutlich daran, dass Noten für Schüler zugängliche Indikatoren (hoher oder niedriger) Fähigkeit sind, Testscores dagegen nicht. Allein die Information über die eigene Leistung beispielsweise in einer Mathematikarbeit reicht aber einem Schüler für den Schluss auf eigene hohe oder niedrige mathematische Fähigkeiten nicht aus. Verschiedene Theorien gehen vielmehr davon aus, dass Personen Vergleichsprozesse anstellen, um einen Schluss von Leistung auf Fähigkeit vorzunehmen (Albert, 1977; Festinger, 1954; Marsh, 1986). Hierbei können unterschiedliche Arten von Leistungsvergleichen unterschieden werden.

Unter sozialen Vergleichen versteht man Vergleiche zwischen verschiedenen Personen hinsichtlich ihrer Leistung. Die Theorie sozialer Vergleichsprozesse (Festinger, 1954) nimmt an, dass Menschen eine zutreffende Wahrnehmung der Höhe ihrer Fähigkeiten anstreben: „The holding of ... inaccurate appraisals of one's abilities can be punishing or even fatal in many situations“ (S. 117; zur Funktionalität zutreffender Fähigkeitsselbstkonzepte vgl. Meyer, 1984). Sofern kein objektives Kriterium zur Fähigkeitsbewertung vorliegt, sollen laut Festinger soziale Vergleiche vorgenommen werden, wobei Vergleiche mit ähnlichen Personen die meisten Informationen für die Fähigkeitsinferenz liefern sollen. Ergibt der soziale Ver-

gleich, dass die Leistungen der Vergleichsgruppe unter den eigenen liegen (sozialer Abwärtsvergleich), so führt dies zu hohen Fähigkeitsselbstkonzepten, während soziale Aufwärtsvergleiche zu niedrigen Fähigkeitsselbstkonzepten führen (vgl. Köller, Klemmert, Möller & Baumert, 1999). Trotz der Funktionalität zutreffender Fähigkeitseinschätzungen postuliert Festinger (1954, p. 124): "There is a unidirectional drive upwards in the case of abilities ...". Jedermann will lieber höhere als niedrigere Fähigkeiten haben („push to do better and better“, p. 125). Hieraus leitet Wagner (1999) ab, dass es eine generelle Präferenz für Aufwärtsvergleiche gibt.¹ Allerdings darf die Diskrepanz zur Vergleichsgruppe nicht zu groß werden, da in diesem Fall der Vergleich keine hilfreichen Informationen liefern würde (Festinger, 1954).

Unter dimensionalen Vergleichen versteht man intraindividuelle Vergleiche, bei denen eigene Leistungen in einer Dimension (bzw. in einem Schulfach) mit eigenen Leistungen in einer weiteren Dimension (bzw. in einem weiteren Schulfach) verglichen werden (z. B. Mathematiknote vs. Deutschnote). Dem internal/external-frame of reference Modell zufolge (Marsh, 1986) nutzen Personen neben sozialen Vergleichen (s.o.) auch dimensionale Vergleiche, um Fähigkeitsselbstkonzepte auszubilden

¹ Ob aus diesem „push“ eindeutig eine Präferenz für Aufwärtsvergleiche ableitbar ist, scheint uns nicht klar. Denkbar wäre beides: Eine Präferenz für Aufwärtsvergleiche, vielleicht weil Aufwärtsvergleiche motivationalen Ansporn bieten oder modellhaft günstigeres Leistungsverhalten veranschaulichen. Denkbar wäre aber auch eine Präferenz für Abwärtsvergleiche, weil sie der Person deutlich machen, dass man das „better and better“ erreicht hat und nun tatsächlich besser als andere ist.

(vgl. hierzu auch Dickhäuser, im Druck). Ergibt der dimensionale Vergleich, dass die Leistungen im Vergleichsfach B unter denen im aktuellen Fach A liegen (dimensionaler Abwärtsvergleich), so führt dies zu einem hohen FSK in A. Analog dazu führen dimensionale Aufwärtsvergleiche zu niedrigen Fähigkeitsselbstkonzepten (vgl. Köller et al., 1999). Diese dimensional Kontrastierungen führen dazu, dass Fähigkeitsselbstkonzepte weitaus geringer (oftmals sogar null-)korreliert sind als die entsprechenden Leistungen. Marsh (1986) berichtet jedoch Befunde, wonach sich bei Schülern aus zweiten und dritten Klassen mit .46 und .49 Korrelationen zwischen mathematischem und verbalem FSK zeigen, die recht nahe an den in der Literatur berichteten Korrelationen der entsprechenden Leistungen liegen. Dies könnte ein Hinweis darauf sein, dass dimensionale Kontrastierungen in diesen Klassenstufen noch keine Effekte haben, wobei unklar ist, ob dies lediglich daran liegt, dass in jüngerem Alter keine dimensional Vergleiche vorgenommen werden.

Unter temporalen Vergleichen versteht man intraindividuelle Vergleiche, bei denen eigene aktuelle Leistungen mit eigenen Leistungen aus der Vergangenheit verglichen werden. Der Theorie temporaler Vergleichsprozesse zufolge (Albert, 1977) haben Personen das Bedürfnis, Selbstwahrnehmungen von unterschiedlichen Zeitpunkten zu einem einheitlichen Selbstbild zu integrieren. Dementsprechend sind Personen motiviert, aktuelle Leistungen mit vergangenen in Beziehung zu setzen. Ergibt der

temporale Vergleich, dass die eigenen früheren Leistungen unter den aktuellen Leistungen liegen (temporaler Abwärtsvergleich), so führt dies zu Wahrnehmungen von Kompetenzzuwachs und damit einem hohen FSK (vgl. hierzu etwa die Befunde von Lüdtke & Köller, 2002). Insbesondere jüngere Schüler/-innen sind geneigt, temporale Vergleiche vorzunehmen (Nicholls, 1984).

Die überwiegende Anzahl an Studien zum Einfluss von Vergleichsprozessen auf Fähigkeitsselbstkonzepte ist pfadanalytischer Natur. Möller und Köller haben 1998 quasiexperimentell untersucht, welche Variablen die Bevorzugung bestimmter Vergleichsprozesse bedingen und welche Effekte die Vergleichsprozesse auf das FSK haben (zu einer experimentellen Analyse der Effekte dimensionaler Vergleiche siehe auch Möller & Köller, 2001). Möller und Köller (1998) befragten Schüler/-innen aus 7. und 8. Klassen eines Gymnasiums kurz vor einer Mathematikarbeit nach dem mathematischen FSK. Unmittelbar nach Rückgabe der Arbeit gaben die Schüler/-innen an, mit welchem anderen Mitschüler (sozialer Vergleich) oder welchem anderen Fach (dimensionaler Vergleich) sie ihre Note in der Klassenarbeit vergleichen möchten. Weiterhin wurde die Richtung des Vergleichs erfasst. Es zeigte sich, dass Schüler mit guten Noten oder hohem FSK vermehrt soziale Horizontal- (Vergleiche mit Schülern mit gleichen Noten) oder Abwärtsvergleiche vornahmen. Schüler mit niedrigem FSK nahmen vermehrt dimensionale Aufwärtsvergleiche vor. Dimensionale Vergleiche haben einen Einfluss auf

das FSK: Nach dimensionalen Aufwärtsvergleichen resultieren niedrigere Fähigkeitsselbstkonzepte als nach dimensionalen Abwärtsvergleichen. Weiterhin zeigten sich nach sozialen Aufwärtsvergleichen niedrigere Fähigkeitsselbstkonzepte als nach sozialen Abwärtsvergleichen, wobei dies nur das Ergebnis der paarweisen Einzelvergleiche ist, während der Effekt des sozialen Vergleichs auf das FSK der Varianzanalyse zufolge die Signifikanzschranke verfehlte.

Entwicklung der Fragestellung und Hypothesen

Die vorliegende Arbeit untersucht leistungsbezogene Vergleichsprozesse und deren Effekte in den Stufen 3 und 4 der Grundschule. Wir wählen mit Grundschulern eine Population, bei der sich das FSK erst zunehmend stabilisiert (vgl. Helmke, 1998) und damit Effekte von Vergleichsprozessen auf Fähigkeitsselbstkonzepte nachweisbar sein sollten. Wegen der Schwierigkeit des Arbeitens mit Fragebögen in den Stufen 1 und 2 beschränken wir die Untersuchung auf Schüler der dritten und vierten Klasse. Die Arbeit versteht sich mit ihrem quasi-experimentellen Design als eine Ergänzung zu bisherigen vornehmlich pfadanalytischen Arbeiten zur Selbstkonzeptentwicklung in der Grundschule (z.B. Faber, 1992).

Zunächst soll durch die vorliegende Arbeit die Präferenz bestimmter Vergleichsrichtungen untersucht werden. In der Arbeit von Möller und Köller (1998) ist nicht überprüft worden, ob die Bevorzugung bestimmter Vergleichsrichtungen durch mangelnde Verfügbarkeit anderer Vergleichsmöglichkeiten erklärt

werden kann: Wählen bestimmte Schüler möglicherweise nur deshalb soziale Horizontal- oder Abwärtsvergleiche, weil sie gar keine Möglichkeit sehen, sich aufwärts zu vergleichen? Die Erfassung der Verfügbarkeit von Vergleichsinformation soll diesen offenen Punkt klären.

Vergleichsprozesse liefern Personen Informationen. Solche Suche nach Informationen sollte insbesondere dann erfolgen, wenn unerwartete Ereignisse eintreten, die den Schemata der Person widersprechen (vgl. Meyer, 1988). Im Falle des Erhalts einer Klassenarbeit ist dies insbesondere dann der Fall, wenn die erhaltene Note von der erwarteten Note abweicht. Dieser Bedeutung der Variable „Erwartungswidrigkeit“ wird in der vorliegenden Studie Rechnung getragen. Wir nehmen an, dass Vergleiche vor allem dem Erhalt einer realistischen Selbsteinschätzung dienen (vgl. Meyer, 1984). Entsprechend vermuten wir, dass sich Schüler nach einer erwartungswidrigen Leistung sozial (Hypothese 1) und temporal (Hypothese 2) in die Richtung vergleichen, die der ursprünglichen Erwartung entspricht. Hinter diesen Hypothesen steht die Annahme, dass diese Vergleiche die meisten Informationen liefern und deshalb bevorzugt werden sollten (vgl. Festinger, 1954; Radloff, 1966; Wood & Taylor, 1991). Goethals und Darley (1977) nehmen an, dass soziale Vergleiche zum Zweck der Fähigkeitsbeurteilung insbesondere mit solchen Personen vorgenommen werden, die vermutlich die gleiche Leistungsfähigkeit haben, wie die Person selbst. Dies trifft auf Personen zu, von denen man annimmt,

dass sie Leistungen erzielen, wie man selbst üblicherweise auch (erwartete Leistung). Erwartet z. B. ein Schüler in einer Arbeit eine 2, erhält aber eine 4, so kann ihm der Vergleich mit denjenigen Mitschülern, von denen er annimmt, dass sie vermutlich ebenfalls eine 2 geschrieben haben, wertvolle Informationen liefern. Stellt sich bei diesem Vergleich nämlich beispielsweise heraus, dass diese Mitschüler ebenfalls eine 4 erhalten haben, liegt nahe, das eigene und fremde schlechte Abschneiden auf eine zu schwere Arbeit zurückzuführen. Diese Argumentation trägt dem Umstand Rechnung, dass es sich beim Schluss von Leistung auf Fähigkeit um einen attributionalen Prozess handelt (Goethals & Darley, 1977).

Hinsichtlich der Effekte der Vergleichsrichtung auf das FSK vermuten wir aufgrund einschlägiger Arbeiten von Collins (1996), dass soziale Aufwärtsvergleiche negative Effekte auf das FSK haben, während sich Abwärtsvergleiche positiv auf das FSK auswirken (Hypothese 3). Diese Hypothese wird gestützt durch Arbeiten zu Bezugsgruppeneffekten, die zeigen, dass lernschwächere Schüler ein niedrigeres FSK aufweisen, wenn sie integriert mit leistungsstärkeren Schülern unterrichtet werden als wenn die Unterrichtung zusammen mit anderen, ebenfalls schwachen Schülern erfolgt (zsf. Wagner, 1999). Die Befunde von Reuman (1989) zeigen, dass dieser Effekt dadurch zustande kommt, dass im Falle einer integrierten Unterrichtung Lernschwache mehr Gelegenheit zu sozialen Aufwärtsvergleichen haben als bei einer separierten Unterrichtung. In der vorliegen-

den Arbeit werden - über die Studie von Möller und Köller (1998) hinausgehend - auch temporale Vergleiche erfasst. Den Vermutungen von Albert (1977) zufolge wäre auch für temporale Vergleiche zu erwarten, dass Aufwärtsvergleiche negative Effekte auf das FSK haben, Abwärtsvergleiche dagegen positive (Hypothese 4a). Im Gegensatz dazu lassen aber die Arbeiten zur individuellen Bezugsnormorientierung von H. Heckhausen (1974) und Rheinberg (1980) vermuten, dass das Vornehmen temporaler Vergleiche eher zur Wahrnehmung von Leistungsvariabilität führt, eher variable Ursachenfaktoren nahe legen (z.B. Anstrengung) und gerade deshalb nicht zu (stabilen) Fähigkeitsrückschlüssen führen sollten. Folgt man diesen Arbeiten, so wäre also konkurrierend zu erwarten, dass temporale Vergleiche keine Veränderung im FSK bewirken (Hypothese 4b).

Die Arbeiten von Marsh (1986, S. 139) lassen vermuten, dass dimensionale Vergleiche bei Drittklässlern noch keine Effekte auf das FSK haben. Im Gegensatz dazu fand Faber (1992) pfadanalytisch Hinweise auf dimensionale Vergleichsprozesse bei Viertklässlern. Für die vorliegende Arbeit mit Dritt- und Viertklässlern erwarten wir daher allenfalls für Viertklässler Effekte dimensionaler Vergleiche auf das FSK.

Methode

Versuchspersonen

An der Untersuchung nahmen insgesamt 195 Schülerinnen und 184 Schüler der Stufen drei (190) und vier (189) aus verschiedenen Grundschulen in Hessen teil. Zum ersten Zeitpunkt betei-

ligten sich 370 Schüler, zum zweiten 361. Von insgesamt 348 Schülern liegen zu den beiden Messzeitpunkten vollständige Datensätze vor. Das Durchschnittsalter der Schüler beträgt 9.6 Jahre. Einverständniserklärungen der Eltern lagen vor.

Untersuchungsablauf und Material

Der erste Teil der Befragung (t1) fand am Tag vor einer Mathematikarbeit statt, der zweite (t2; im Schnitt zwei Tage später) direkt nach der Rückgabe der Mathematikarbeit. Zu t1 wurde das FSK erfasst. Zu t2 wurden die Vergleichsprozesse und noch einmal das FSK erfasst. Die Schüler/-innen machten zunächst demografische Angaben zu Geschlecht, Alter und Klasse. Das FSK wurde zu t1 und zu t2 in Anlehnung an Tiedemann (2000) wie folgt erfasst: „Ich bin gut in Mathe“, „Mathe fällt mir leicht“ und „Oft verstehe ich Mathe einfach nicht“. Die Items wurden auf einer Skala von „stimmt gar nicht“ [1] bis „stimmt genau“ [5] beantwortet. Es wurde auf Items verzichtet, die den Vergleich mit Klassenkameraden thematisieren. Die Verwendung von Items mit ausdrücklich fähigkeitsthematischem Inhalt wurde unterlassen, da sich in vorangegangenen Untersuchungen in dritten Klassen herausgestellt hatte, dass einige Schüler bei dem Begriff „Fähigkeit“ Verständnisschwierigkeiten hatten. Die drei Items bilden zu t1 und zu t2 (bei Recodierung des dritten Items) eine befriedigend homogene Skala (Cronbach's α jeweils .81). Zu t1 wurde zusätzlich die Notenerwartung wie folgt erfasst: „Was glaubst du, welche Note du in der nächsten Mathematikarbeit schreiben wirst?“

Zu t2 wurde den Schülern nach der Aushändigung der Fragebögen ihre benotete Arbeit zurückgegeben. Die Schüler schauten in ihre Hefte und bearbeiteten anschließend sofort den Fragebogen. Sie wurden angewiesen, sich während der Befragung untereinander nicht über die erhaltenen Noten auszutauschen und sich möglichst nicht anmerken zu lassen, ob die erhaltene Note gut oder schlecht sei. Das FSK wurde wie zu t1 erfasst. Die Leistung der Schüler in der Mathematikarbeit wurde wie folgt erfasst: „Welche Note hast du gerade in der Mathearbeit bekommen?“. (Die Angaben zu dieser Frage wurden mit den Lehrerangaben zur Note abgeglichen. Es gab keine Abweichungen.) Die Verfügbarkeit des sozialen / dimensionalen / temporalen Aufwärts- bzw. Abwärtsvergleichs wurde folgendermaßen erfasst: „Glaubst du, es gibt in deiner Klasse einen Mitschüler, der in der Mathearbeit besser (schlechter) ist als du?“ (sozial) bzw. „Glaubst du, es gibt ein Fach in deinem letzten Zeugnis, in dem du besser (schlechter) als in der Mathearbeit bist?“ (dimensional) und „Glaubst du, dass du früher in Mathe mal besser (schlechter) warst als heute in der Mathearbeit?“ (temporal). Es wurden jeweils die Antwortmöglichkeiten angeboten: „Ja, ich glaube, dass ...“ bzw. „Nein, ich glaube nicht, dass ...“.

Die Erfassung der Vergleichsrichtung sozial / dimensional / temporal wurde durch die frei zu beantwortende Frage nach dem Subjekt, Objekt bzw. Zeitpunkt des Vergleichs eingeleitet (sozial: „Wenn du deine Note in der Mathearbeit mit der Note von einem Mitschüler aus deiner Klasse vergleichst, an welchen

Mitschüler denkst du dann?"; dimensional: „...mit der Note in einem anderen Fach aus deinem letzten Zeugnis vergleichst, an welches Fach denkst du dann?"; temporal: „... mit einer Note vergleichst, die du früher in Mathe hattest – wann [z. B. in welcher Klasse, vor wie vielen Wochen/Monaten] hattest du diese Note?“).

Die Richtung der Vergleiche wurde folgendermaßen erfasst: „Glaubst du, dieser Mitschüler ist in der Mathearbeit besser als du, genauso gut wie du oder schlechter als du?“ (sozial) bzw. „Glaubst du, du bist in dem anderen Fach besser als in der Mathearbeit, genauso gut wie in der Mathearbeit oder schlechter als in der Mathearbeit?“ (dimensional) und „Glaubst du, du warst damals besser als heute in der Mathearbeit, genauso gut wie heute in der Mathearbeit oder schlechter als heute in der Mathearbeit?“ (temporal). Als Antwortmöglichkeiten standen den Schülern jeweils die Möglichkeiten „besser“, „genauso gut“ und „schlechter“ zur Verfügung. Der Wunsch nach dem Vornehmen der Vergleiche wurde über folgende Items erhoben: „Möchtest du deine Note in der Mathearbeit überhaupt mit der Note von einem Mitschüler (sozial) / mit der Note in einem anderen Fach (dimensional) / damit wie du früher in Mathe warst (temporal) vergleichen?“ Als Antwortmöglichkeiten standen jeweils „Ja, ich möchte ...“ bzw. „Nein, ich möchte meine Note nicht...“ zur Verfügung. Die Befragung dauerte zu t1 ungefähr 20 und zu t2 circa 45 Minuten. Die Fragebögen wurden

von einer geschulten Versuchsleiterin in standardisierter Form vorgegeben und erläutert.

Ergebnisse

Zuerst wurde betrachtet, wie viele Schüler sich bei der jeweiligen Vergleichsart abwärts, horizontal und aufwärts vergleichen (vgl. Tabelle 1). Es zeigte sich beim sozialen Vergleich, dass die Vergleichsrichtungen nicht in gleicher Weise bevorzugt werden, $\chi^2 (2, \underline{N} = 360) = 72.27, \underline{p} < .001$. Soziale Abwärtsvergleiche werden nur von 12.2 Prozent der Schüler vorgenommen und sind damit seltener als Horizontal- und Aufwärtsvergleiche. Betrachtet man die Frage nach der von den Schülern angegebenen Verfügbarkeit von sozialen Abwärtsvergleichen, so stehen diesen 12.2 Prozent, die sich abwärts vergleichen, 91.3 Prozent an Schülern gegenüber, die angeben, über die Möglichkeit zum sozialen Abwärtsvergleich zu verfügen. Auch dimensionale und temporale Vergleiche werden nicht in alle Richtungen gleich häufig vorgenommen, dimensional: $\chi^2 (2, \underline{N} = 359) = 120.07, \underline{p} < .001$, temporal: $\chi^2 (2, \underline{N} = 359) = 137.66, \underline{p} < .001$. Dimensionale und temporale Abwärtsvergleiche sind am seltensten (10.9 bzw. 16.4 Prozent), dimensionale und temporale Aufwärtsvergleiche am häufigsten. Den 10.9 Prozent an Schülern, die sich dimensional abwärts vergleichen stehen 45 Prozent an Schülern mit einer entsprechenden Möglichkeit gegenüber. Beim temporalen Vergleich stehen den 16.4 Prozent 40.8 Prozent mit der entsprechenden Möglichkeit gegenüber.

Um die Vergleichsprozesse in Abhängigkeit von der Erwartungswidrigkeit der Leistung analysieren zu können, wurden die Schüler aufgrund des Vergleichs zwischen erwarteter und erhaltener Note in Gruppen eingeordnet. Schüler, welche die Note 1 ($\underline{n} = 61$) bzw. die Note 5 oder 6 ($\underline{n} = 9$) erwarten, wurden ausgeschlossen, da im ersten Fall ein Übertreffen, im zweiten Fall ein Untertreffen der Erwartung nicht möglich ist. Es verblieben in der Gruppe „Erwartung untertroffen“ 133, in der Gruppe „Erwartung entsprochen“ 109 und in der Gruppe „Erwartung übertroffen“ 43 Schüler. Es wurden dann für jede der Vergleichsarten die Richtung des Vergleichs und die Erwartungswidrigkeit kreuztabelliert. In Tabelle 2 sind die beobachteten und (in Klammern) die erwarteten Häufigkeiten getrennt für soziale Vergleiche (oberes Drittel), dimensionale Vergleiche (mittleres Drittel) und temporale Vergleiche (unteres Drittel) dargestellt.

Hinsichtlich der sozialen Vergleiche werden die Vergleichsrichtungen je nach Erwartungswidrigkeit unterschiedlich bevorzugt, $\chi^2 (4; \underline{N} = 285) = 18.18$, $\underline{p} < .002$, $\Phi = .25$. Untertrefft die Leistung die Erwartung, vergleichen sich mehr Schüler als erwartet aufwärts (also in Richtung der erwarteten Note) und weniger Schüler als erwartet horizontal. Erfüllt die Leistung die Erwartung, vergleichen sich mehr Schüler als erwartet in Richtung der Erwartung (horizontal), während sich weniger Schüler als erwartet aufwärts vergleichen. Ist die

Leistung besser als erwartet, sind ebenfalls bei mehr Schülern als erwartet Vergleiche in Richtung der ursprünglichen Notenerwartung zu verzeichnen. Weiterhin vergleichen sich mehr dieser Schüler als erwartet horizontal und weniger als erwartet aufwärts.

Auch der dimensionale Vergleich wird je nach Erwartungswidrigkeit in unterschiedliche Richtungen bevorzugt, $\chi^2 (4; N = 285) = 53.27, p < .001, \Phi = .43$. Wird die Erwartung untertroffen, folgt häufiger als erwartet ein Vergleich in Richtung der ursprünglichen Notenerwartung (dimensionaler Aufwärtsvergleich). Horizontal- und Abwärtsvergleiche werden nach untertroffener Erwartung von weniger Schülern als erwartet angewendet. Bei entsprechender Erwartung vergleichen sich Schüler häufiger als erwartet horizontal, während seltener als erwartet ein Aufwärtsvergleich durchgeführt wird. Bei übertroffener Erwartung werden Abwärts- und Horizontalvergleiche häufiger als erwartet vorgenommen. Aufwärtsvergleiche kommen seltener vor als erwartet.

Schließlich wird auch der temporale Vergleich je nach Erwartungswidrigkeit in unterschiedliche Richtungen bevorzugt, $\chi^2 (4; N = 285) = 53.27, p < .001, \Phi = .50$. Nach untertroffener Erwartung führen Schüler häufiger als erwartet Aufwärtsvergleiche durch, während Horizontal- und Abwärtsvergleiche seltener als erwartet vorgenommen werden. Entspricht die Note der Erwartung, vergleichen sich mehr Schüler als erwartet ho-

horizontal. Aufwärtsvergleiche werden dagegen dann seltener als erwartet durchgeführt. Übertrifft die Leistung die Erwartung, vergleichen sich mehr Schüler als erwartet abwärts und horizontal. Den Aufwärtsvergleich nehmen weniger Schüler als erwartet vor.

Zur Analyse der Effekte der Vergleichsrichtungen auf das FSK zu t2 wurden getrennt für die Vergleichsarten sozial, dimensional und temporal univariate Kovarianzanalysen berechnet. Es wurde geprüft, ob der Faktor „Vergleichsrichtung“ („aufwärts“, „horizontal“, „abwärts“) einen Effekt auf die mittleren Werte des FSK (t2) hat. Das FSK zu t1 wurde als Kovariate eingeführt.

Es ist zu vermuten, dass die Vergleichsprozesse nur bei solchen Kindern Auswirkungen auf das FSK haben, die ausdrücklich einen solchen Vergleich auch durchführen wollen. Es wurden daher nur die Personen in der Analyse berücksichtigt, die angegeben hatten, sich sozial ($n = 287$) bzw. dimensional ($n = 239$) oder temporal ($n = 245$) überhaupt vergleichen zu wollen.

Für den Faktor „Vergleichsrichtung sozial“ erbringt die Kovarianzanalyse zunächst einen signifikanten Effekt der Kovariaten auf das FSK zu t2, $F(1; 283) = 653.24$, $p < .001$, $\eta^2 = .70$. Zudem ergibt sich ein signifikanter Effekt des Faktors „Vergleichsrichtung sozial“ auf das FSK zu t2, $F(2; 283) = 4.97$, $p < .01$, $\eta^2 = .03$. Angeschlossene Mittelwertsvergleiche nach Bonferroni unter Berücksichtigung der Kovariaten ergeben,

dass das FSK zu t2 nach Aufwärtsvergleichen signifikant niedriger als nach Abwärtsvergleichen ist (vgl. Tabelle 3).²

Hinsichtlich der unabhängigen Variable „Vergleichsrichtung dimensional“ zeigt sich in der Kovarianzanalyse ebenfalls ein signifikanter Effekt der Kovariaten FSK (t1) ($F(1; 235) = 524.53$, $p < .001$, $\eta^2 = .70$). Allerdings ergibt die Analyse des Faktors „Vergleichsrichtung dimensional“ keinen signifikanten Effekt auf das FSK. Rechnet man diese Analyse unter Verwendung des zusätzlichen Faktors Schulstufe, so zeigen sich keine Interaktionen von Schulstufe und Vergleichsrichtung.

Bei der Variable „Vergleichsrichtung temporal“ ergibt sich ein signifikanter Effekt der Kovariaten ($F(1; 241) = 535.60$, $p < .001$, $\eta^2 = .70$), aber kein Effekt des Faktors „Vergleichsrichtung temporal“ (alle F -Werte < 2.65).

Diskussion

² Um diese Effekte der Vergleichsrichtung eindeutig als solche interpretieren zu können, weist ein anonymer Gutachter dankenswerterweise darauf hin, dass die Erwartungswidrigkeit (welche ja mit der Vergleichsrichtung zusammenhängt) als Kovariate eingeführt werden soll. In dieser ersten Kovarianzanalyse kann man aber nur die Versuchspersonen mit einbeziehen, deren Notenewartung nicht 1 oder 5/6 ist, da bei diesen die Erwartungswidrigkeit in Teilen Folge der Erwartung ist. Dadurch reduziert man die Anzahl der Versuchspersonen. Zur Beurteilung der dann vorhandenen Effekte im Vergleich zu ersten Analyse sollte nun (wegen des reduzierten n) nicht auf die Signifikanz zurückgegriffen werden. Wir geben daher stattdessen für die Kovariaten und den Faktor jeweils die Effektstärken an: Kovariate FSK (t1): $\eta^2 = .66$, Kovariate Erwartungswidrigkeit: $\eta^2 = .05$, Faktor Vergleichsrichtung: $\eta^2 = .03$. In einem letzten Schritt wurde, den Wünschen eines zweiten Gutachters folgend die Note als Kovariate mit einbezogen. In dieser Analyse wiederum müssen zusätzlich die Schüler mit den erhaltenen Noten 1 bzw. 5/6 zusätzlich ausgeschlossen werden, da erstere keine Möglichkeit zum Aufwärts-, letztere keine Möglichkeit zum Abwärtsvergleich haben. Die Effektstärken waren wie folgt: Kovariaten FSK (t1): $\eta^2 = .49$, Kovariate Erwartungswidrigkeit: $\eta^2 = .01$, Kovariate Note: $\eta^2 = .08$, Faktor Vergleichsrichtung: $\eta^2 = .03$. In allen drei Analysen bleibt der Effekt der Vergleichsrichtung auf die Selbstkonzeptveränderung bei $\eta^2 = .03$.

In der vorliegenden Arbeit wurden leistungsbezogene Vergleichsprozesse im Zusammenhang mit dem FSK in der Grundschule untersucht. Im Unterschied zu bisherigen Studien (Möller, 2000; Möller & Köller, 1998) wurde in der vorliegenden Studie die Erwartungswidrigkeit der Leistungen berücksichtigt. Anders als in bisherigen Arbeiten (z. B. Helmke, 1991, 1998) haben wir das FSK durch solche Items erfasst, in denen ausdrücklich auf die Thematisierung sozialer Vergleiche verzichtet wird. Erst ein solches Vorgehen erlaubt einen klaren Nachweis der Effekte der Vergleichsprozesse auf das FSK.

Schließlich untersuchten wir in der vorliegenden Untersuchung Vergleichsprozesse in einer realen Rückmeldesituation, womit die Befunde valide Auskunft darüber versprechen, welche Prozesse zur Veränderung von Fähigkeitsselbstkonzepten führen.

Es zeigte sich, dass bei allen Vergleichsarten Aufwärtsvergleiche (also Vergleiche mit besseren Mitschülern, besseren Fächern oder Zeitpunkten, zu denen man besser war) gegenüber Abwärtsvergleichen bevorzugt werden. Das seltene Vornehmen von Abwärtsvergleichen kann, wie sich zeigte, nicht (oder zumindest nicht ausschließlich) durch unzureichende Verfügbarkeit von Vergleichsmöglichkeiten erklärt werden. Die häufige Nutzung von Aufwärtsvergleichen ist möglicherweise dadurch zu erklären, dass diese informativer sind als Abwärtsvergleiche: Der Vergleich mit besseren Schülern, Fächern oder Zeitpunkten, zu denen man besser war, liefert nämlich Beispiele für erfolgreiches Leistungsverhalten. Dies gilt für soziale Vergleichs-

prozesse (Seta, 1982), aber auch dimensionale und temporale Aufwärtsvergleiche geben wichtige Hinweise darüber, welche Strategien in anderen Fächern oder zu anderen Zeitpunkten zu besseren Leistungen geführt haben als im Referenzfach.

Die Bevorzugung von Aufwärts-, Horizontal- oder Abwärtsvergleichen erfolgt jedoch in Abhängigkeit von der Erwartungswidrigkeit der Leistung unterschiedlich (vgl. Hypothese 1 und 2). Wie erwartet zeigten sich für soziale und temporale Vergleiche (ebenso wie für dimensionale Vergleiche) häufiger Vergleiche in die Richtung, die der ursprünglichen Erwartung entsprechen (Aufwärtsvergleiche nach unertroffener, Horizontalvergleiche nach entsprochener und Abwärtsvergleiche nach übertrroffener Erwartung). Zu vermuten ist, dass diese Vergleiche den Schülern dazu dienen, zu einer realistischen Einschätzung der Höhe eigener Fähigkeiten zu gelangen, denn um die eigenen Fähigkeiten aus der Leistung erschließen zu können, sollten Vergleiche in Richtung der ursprünglichen Erwartung am informativsten sein.

Im Falle einer erwartungswidrig positiven Leistung ergibt sich für alle Vergleicharten auch eine häufige Nutzung von Horizontalvergleichen. Der Zusammenhang zwischen Erwartungswidrigkeit und Vergleichsprozessen ist möglicherweise nicht symmetrisch. Der Grund für die Bevorzugung des sozialen Horizontalvergleichs nach einer erwartungswidrig positiven Leistung könnte darin liegen, dass sich diese Schüler/-innen bereits schneller als der Gruppe derjenigen mit einer guten Leistung

zugehörig bzw. ihr ähnlich fühlen. Dieses Zugehörigkeitsgefühl wiederum kann sich positiv auf das FSK auswirken.

Ein anonymes Gutachter merkt an, die Variable „Erwartungswidrigkeit“ sei mit der ursprünglichen Erwartung konfundiert. Für einen Schüler, der die Note „2“ erwartet, stehen auf der Notenskala mehr Notenstufen unterhalb als oberhalb seiner Erwartung zur Verfügung (analog bei der Notenerwartung „4“ mehr oberhalb als unterhalb). Wenn man gewisse Zufallseinflüsse auf die Notengebung annimmt, ist hieraus ableitbar, dass allein aufgrund der Erwartung „2“ ein Untertreffen etwas wahrscheinlicher ist als ein Übertreffen und aufgrund der Erwartung „4“ ein Übertreffen etwas wahrscheinlicher als ein Untertreffen. Dieses Problem ist auch nicht dadurch komplett lösbar, dass man nur Schüler mit der Notenerwartung „3“ in die Analyse mit einbezieht, denn auch hier gibt es mit den Notenstufen 4, 5 und 6 mehr potentielle Notenstufen unterhalb als oberhalb der ursprünglichen Erwartung (1 und 2). Insofern muss – allerdings nur wenn man Zufallseinflüsse bei der Notenvergabe annimmt – bei den vorliegenden Ergebnissen offen bleiben, ob die Effekte der Erwartungswidrigkeit in gewissen Teilen auch als Effekte der ursprünglichen Erwartung zu erklären sind.

Es zeigte sich, dass der überwiegende Anteil an Schülern angibt, sowohl soziale, dimensionale wie auch temporale Vergleiche vornehmen zu wollen (wobei soziale Vergleichsinformation von den meisten Schülern gewünscht ist). Dies entspricht

den bisherigen Annahmen aus der Literatur, denen zufolge Schüler dieses Alters vermehrt soziale Vergleiche (Helmke, 1998) aber auch temporale Vergleiche (Nicholls, 1984) vornehmen.

Die Ergebnisse hinsichtlich des Einflusses sozialer Vergleiche auf das FSK stimmen mit der Hypothese 3 überein: Soziale Abwärtsvergleiche wirken sich positiv auf das FSK aus, während Aufwärtsvergleiche negative Folgen für das FSK haben. Die vorliegende Arbeit liefert damit einen quasiexperimentellen Beleg dafür, dass Leistungsrückmeldungen in der Grundschule vermittelt über soziale Vergleichsprozesse zu Veränderungen im FSK führen können. Der Befund hinsichtlich der Effekte sozialer Vergleiche ist klarer als die nur tendenziellen Veränderungen in der Arbeit von Möller und Köller (1998). Die Höhe der Effektstärke ist als schwach zu bezeichnen, allerdings kann man bei einem so kurzen Zeitabstand bei der Erfassung der FSK kaum stärkere Einflüsse erwarten.

Trotz der positiven Auswirkungen von sozialen Abwärtsvergleichen auf das FSK werden diese wie bereits beschrieben bei Dritt- und Viertklässlern vergleichsweise selten genutzt. Stattdessen kommen FSK-schädliche Aufwärtsvergleiche verstärkt zum Einsatz. Möglicherweise nutzen die Schüler die sozialen Vergleiche nicht in erster Linie, um ihr FSK zu erhöhen (Selbstaufwertungsansatz, vgl. z.B. J. Heckhausen & Krüger, 1993), sondern die durch den Aufwärtsvergleich gebotenen Möglichkeiten zum Modelllernen führen zur Bevorzugung dieser Vergleichsrichtung. Untersuchungen, in denen auch die Vergleichs-

motive der Kinder erfasst werden, können hier Aufschluss geben.

Für den dimensionalen Vergleich ergeben sich keine Effekte auf das FSK - wie die vorliegende Arbeit für die Grundschule erstmals quasiexperimentell zeigt. Entgegen den Befunden bei Jugendlichen und Erwachsenen (Möller & Köller, 1998; Möller, 2000) beeinflussen dimensionale Vergleiche weder bei Dritt- noch bei Viertklässlern in der vorliegenden Untersuchung das FSK. Das Ausbleiben des Effekts kann gemäß unserer Analyse nicht dadurch erklärt werden, dass Schüler in der Grundschule gar keine dimensionalen Vergleiche vornehmen wollen. Der Grund für das Ausbleiben des Effekts könnte darin liegen, dass die verschiedenen Fächer in der Grundschule von ein und derselben Lehrerin unterrichtet werden, womit selbst unterschiedliche Leistungen in verschiedenen Fächern möglicherweise den Schülern als wenig distinkt erscheinen. Eine frühere Untersuchung von Faber (1992) fand negative Koeffizienten für den Pfad von verbaler Leistung auf das mathematische FSK bei Viertklässlern. Dies wurde vom Autor als ein Hinweis auf dimensionale Vergleiche der mathematischen mit der verbalen Leistung interpretiert. In unserer Untersuchung waren die Schüler jedoch frei, das Vergleichsfach zu wählen. Anhand dieses die Schüler beim Vergleich weniger stark einschränkenden Vorgehens bleiben Effekte des dimensionalen Vergleichs auf das Selbstkonzept aus.

Gemäß den Ergebnissen der vorliegenden Untersuchung haben auch temporale Vergleiche in der Grundschule keine signifikante Bedeutung bei der Formierung des FSK. Dieser Befund steht im Gegensatz zu der Hypothese 4a und den Ergebnissen anderer Untersuchungen, die für einen Einfluss temporaler Vergleiche im Grundschulalter sprechen (Butler, 1998; Ruble, Eisenberg, & Higgins, 1994; Ruble & Frey, 1991; Suls & Mullen, 1982). Es zeigte sich also wie in der Alternativhypothese 4b vermutet in Abhängigkeit vom temporalen Vergleich keine differentielle Veränderung des FSK. Eine mögliche Erklärung hierfür liefern die Überlegungen von H. Heckhausen (1974) und Rheinberg (1980), wonach temporale Vergleiche (individuelle Bezugsnorm) variable Ursachenfaktoren zur Erklärung von Leistungsunterschieden nahe legen und damit stabile Fähigkeitsattributionen wenig wahrscheinlich sind. Eine andere mögliche Erklärung liefern die Überlegungen Alberts (1977), der annimmt, dass temporale Vergleiche nur dann Einfluss auf das FSK haben, wenn weder objektive noch soziale Vergleiche durchgeführt werden können. Da in der vorliegenden Untersuchung soziale Vergleiche durchgeführt wurden, könnte dies zur Folge haben, dass temporale Vergleiche unter dieser Bedingung keinen zusätzlichen Einfluss auf das FSK haben.

Wir haben diskutiert, dass Vergleichsprozesse verschiedenen Zielen dienen können, beispielsweise der Selbstbewertung, der Selbstaufwertung oder dem Modelllernen. Entsprechend ist vorstellbar, dass die Effekte von leistungsbezogenen Vergleichen

chen vielfältig sind und sich nicht nur auf das Fähigkeitsselbstkonzept beziehen. Entsprechende Untersuchungen, in denen etwa -- wie aufgrund des Modelllernens zu vermuten -- positive Auswirkungen von sozialen Aufwärtsvergleichen auf nachfolgendes Lernverhalten nachgewiesen werden, stehen bislang noch aus.

Literaturverzeichnis

Albert, S. (1977). Temporal comparison theory. Psychological Review, 84, 485-503.

Aken, M. A. G. van, Helmke, A. & Schneider, W. (1997). Selbstkonzept und Leistung - Dynamik ihres Zusammenspiels: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. Weinert & A. Helmke (Hrsg.), Entwicklung im Grundschulalter (S. 341-350). Weinheim: Beltz.

Butler, R. (1998). Age trends in the use of social and temporal comparison for self-evaluation: Examination of a novel developmental hypothesis. Child Development, 69, 1054-1073.

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. Psychological Bulletin, 119, 51-69.

Dickhäuser, O. (im Druck). Überprüfung des erweiterten Modells des internal/external frame of reference. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie.

Faber, G. (1992). Bereichsspezifische Beziehungen zwischen leistungsthematischen Schülerselbstkonzepten und Schulleistungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 24, 66-82.

Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7, 117-140.

Goethals, G. R. & Darley, J. M. (1977). Social comparison theory. An attributional approach. In J. M. Suls & R. L. Miller (Eds.), Social comparison process. Theoretical and empiri-

cal perspectives (pp. 259-278). Washington, DC: Hemisphere publishing corporation.

Heckhausen, H. (1974). Leistung und Chancengleichheit. Göttingen: Hogrefe.

Heckhausen, J. & Krüger, J. (1993). Developmental expectations for the self and most other people: Age grading in three functions of social comparison. Developmental Psychology, 29, 539-548.

Helmke, A. (1991). Entwicklung des Fähigkeitsselbstbildes vom Kindergarten bis zur dritten Klasse. In R. Pekrun & H. Fend (Hrsg.), Schule und Persönlichkeitsentwicklung (S. 83-99). Stuttgart: Enke.

Helmke, A. (1992). Selbstvertrauen und schulische Leistungen. Göttingen: Hogrefe.

Helmke, A. (1998). Vom Optimisten zum Realisten? Entwicklung des Fähigkeitsselbstkonzeptes vom Kindergarten bis zur 6. Klassenstufe. In F. E. Weinert (Hrsg.), Entwicklung im Kindesalter (S. 115-132). Weinheim: Beltz.

Helmke, A. & van Aken, M. A. G. (1995). The causal ordering of academic achievement and self-concept during elementary school: A longitudinal study. Journal of Educational Psychology, 87, 624-637.

Köller, O., Klemmert, H., Möller, J. & Baumert, J. (1999). Eine längsschnittliche Überprüfung des Modells des internal/external frame of reference. Zeitschrift für Pädagogische Psychologie, 13, 128-134.

Lüdtke, O. & Köller, O. (2002). Individuelle Bezugsnormorientierung und soziale Vergleiche im Mathematikunterricht: Einfluss unterschiedlicher Referenzrahmen auf das fachspezifische Selbstkonzept der Begabung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 34, 156-166.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal / external frame of reference model. American Educational Research Journal, 23, 129-149.

Marsh, H. W. & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. Educational Psychologist, 20, 107-123.

Meyer, W.-U. (1984). Das Konzept von der eigenen Begabung. Bern: Hans Huber.

Meyer, W.-U. (1988). Die Rolle von Überraschung im Attributionsprozess. Psychologische Rundschau, 39, 136-147.

Möller, J. (2000). Effekte dimensionaler und sozialer Vergleiche auf Fähigkeitseinschätzungen und die Zufriedenheit mit der Leistung. Zeitschrift für Experimentelle Psychologie, 47, 67-71.

Möller, J. & Köller, O. (1998). Dimensionale und soziale Vergleiche nach schulischen Leistungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 30, 118-127.

Möller, J. & Köller, O. (2001). Dimensional comparison: An experimental approach to the internal/external frame of reference model. Journal of Educational Psychology, 93, 826-835.

Nicholls, J. G. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. Psychological Review, *91*, 328-346.

Radloff, R. (1966). Social comparison and ability evaluation. Journal of Experimental Social Psychology Supplement, *1*, 6-26.

Reuman, D. A. (1989). How social comparison mediates the relation between ability-grouping practices and students' achievement expectancies in mathematics. Journal of Educational Psychology, *81*, 178-189.

Rheinberg, F. (1980). Leistungsbewertung und Lernmotivation. Göttingen: Hogrefe.

Ruble, D. N., Eisenberg, R. & Higgins, E. T. (1994). Developmental changes in achievement evaluation: Motivational implications of self-other differences. Child Development, *65*, 1095-1110.

Ruble, D. N. & Frey, K. S. (1991). Changing patterns of comparative behavior as skills are acquired: A functional model of self-evaluation. In J. Suls & T. A. Wills (Eds.), Social comparison: Contemporary theory and research (pp. 79-113). Hillsdale, NJ: Erlbaum.

Seta, J. (1982). The impact of comparison processes on coactors' task performance. Journal of Personality and Social Psychology, *42*, 281-291.

Suls, J. M. & Mullen, B. (1982). From the cradle to the grave: Comparison and self-evaluation across the life-span. In

J. M. Suls (Ed.), Psychological Perspectives on the Self (Vol. 1, pp. 97-125). Hillsdale, NJ: Erlbaum.

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of childrens' concept of their mathematical ability in elementary school. Journal of Educational Psychology, 92, 1-8.

Wagner, J. W. L. (1999). Soziale Vergleiche und Selbsteinschätzungen. Waxmann: Münster.

Wigfield, A. & Karpathian, M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. Educational Psychologist, 26, 233-261.

Wood, J. & Taylor, K. L. (1991). Serving self-relevant goals through social comparison. In J. Suls & T. A. Wills (Eds.), Social comparison: Contemporary theory and research (pp. 23-51). Hillsdale, NJ: Erlbaum.

Tabelle 1

Prozentuale Anteile der Schüler, die sich abwärts, horizontal und aufwärts vergleichen (Angaben in Zeilenprozent)

Vergleichsart	Vergleichsrichtung		
	abwärts	horizontal	aufwärts
sozial	12.2	44.4	43.4
dimensional	10.9	31.2	57.9
temporal	16.4	21.2	62.4

Tabelle 2

Präferenz von Vergleichsrichtungen in Abhängigkeit von der Erwartungswidrigkeit der Leistung. Angaben getrennt für soziale (oberes Tabellendrittel), dimensionale (mittleres Tabellendrittel) und temporale (unteres Tabellendrittel) Vergleiche. Dargestellt sind in jeder Zelle die beobachteten sowie (in Klammern) die erwarteten Häufigkeiten

	Erwartung untertroffen	Erwartung entsprochen	Erwartung übertroffen
sozial			
aufwärts	76 (60)	42 (49)	11 (20)
horizontal	43 (58)	57 (48)	25 (19)
abwärts	14 (15)	10 (12)	7 (4)
dimensional			
aufwärts	107 (78)	47 (64)	13 (25)
horizontal	20 (41)	48 (34)	20 (13)
abwärts	6 (14)	14 (11)	10 (5)
temporal			
aufwärts	112 (84)	57 (69)	11 (27)
horizontal	7 (26)	36 (21)	12 (8)
abwärts	14 (23)	16 (19)	20 (8)

Tabelle 3

Mittleres FSK zu t2 unter Berücksichtigung der Kovariaten in
Abhängigkeit von der Richtung sozialer Vergleiche (Standard-
fehler in Klammern)

Vergleichsart	Vergleichsrichtung			p-Wert
	aufwärts	horizontal	abwärts	
sozial	3.68 ^a (.05)	3.84 (.05)	3.98 ^b (.09)	< .01

Anmerkungen.

Die angegebenen Werte sind das Ergebnis univariater Kovarianzanalysen. Innerhalb einer Zeile kennzeichnen unterschiedliche Indizes signifikant voneinander verschiedene Werte.