CURRENT RESEARCH IN SOCIAL PSYCHOLOGY

http://www.uiowa.edu/~grpproc/crisp/crisp.html

Volume 10, Number 12
Submitted: November 4, 2004
First Revision: January 3, 2005
Second Revision: March 13, 2005
Accepted: March 15, 2005
Published: March 15, 2005

SOCIAL COMPARISON IN THE CLASSROOM: IS THERE A TENDENCY TO COMPARE UPWARD IN ELEMENTARY SCHOOL?

Florence Dumas, Pascal Huguet, Jean-Marc Monteil
Centre National de la Recherche Scientifique (CNRS), Université Aix-Marseille I, France

Claire Rastoul
Université Blaise Pascal, Clermont-Ferrand, France

John B. Nezlek
College of William and Mary, Williamsbourg, USA

ABSTRACT

The central question here is whether elementary school children compare their exam grades with other children in their classroom who perform slightly better than themselves, as typically do middle school children. Children in grade levels five through nine nominated their comparison targets in three academic domains, and a series of standard regression and multilevel analyses examined the relationships between children's performances and the performances of their targets in these domains. Children in grade levels five and six did not compare upward, whereas children in grade levels seven and eight did in some courses, and children in ninth grade level did in each course. The present results clearly demonstrate that the tendency to compare upward becomes stronger over time in the school system.

INTRODUCTION

Although a number of investigators have recognized the importance of social comparison in educational settings (e.g., Levine, 1983; Ruble & Frey, 1991), students' comparison-level choice (i.e., the level typical of the persons with whom they choose to compare) has not received much attention. In a recent study, however, Blanton, Buunk, Gibbons, and Kuyper (1999) found that students compared their exam grades with other (same-sex) students in their classroom who performed slightly better than themselves, exactly as Festinger (1954) would predict. In this study, Dutch middle school children (ninth grade level) listed on a questionnaire their usual comparison target in each of seven courses at the end of trimester 2 (T2). The participants did not estimate their own exam grade nor did they estimate their comparison-level; both of these scores were taken from official grade records and so the relationship between these two variables could not be due to a self-report bias. A score of comparison-level choice was then assigned to each student on the basis of the course grade their comparison-others received at the time of the nomination, and the associations between this choice and their own course grades at T2 were tested. Not only did students compare slightly upward with their classmates (as revealed by paired t-tests between students' trimester grades in the different courses and those of their comparison targets), but higher course grades were associated with higher comparison-level choices (as revealed by standard regression analyses).

As noted by Blanton et al. (1999), the reasons why students might engage in upward comparisons with their classmates are numerous. First, observing another person who has proficiency at a task can reveal useful information about how to improve (e.g., Buunk & Ybema, 1997). Second, seeing another person succeed may increase the motivation to improve. But this latter reason is not as straightforward as the first, because it may be motivating to see others doing well at a task for a variety of reasons. Individuals may come to identify with successful targets, leading to imitation of the targets' actions (e.g., Bandura, 1986). Viewing others succeed may also lead individuals to set higher personal standards for evaluating their own success, which can motivate efforts toward these new and more challenging goals (Seta, 1982). Finally, observing others doing well can endow individuals with a sense of their own potential (e.g., Buunk, Collins, Taylor, Van Yperen, & Dakof, 1990), and this can raise self-confidence and feelings of self-efficacy at the task. There is indeed ample evidence that these feelings play a significant role in academic achievement (e.g., Zimmerman, 1995).

Blanton et al. (1999) captured students' self-efficacy by measuring their comparative evaluation (i.e., how they evaluated their standing in the different courses compared to most of their classmates), which may not reflect social comparison processes per se (Wood, 1996). Consistent with this, choosing to compare with someone who outperformed them in a course did not leave participants feeling relatively less able in that course. When they made their comparative evaluations, participants reflected more on their own abilities than on the performances of their comparison targets.

Huguet, Dumas, Monteil, and Genestoux (2001) replicated and extended these findings (with children of the same age) in a number of important ways. First of all, they offered a more detailed record of comparison choices. In Blanton et al.'s study, participants were asked to nominate only one student with whom they typically compared their exam grades in each course. As pointed out by Blanton et al. themselves, however, students who were interested in obtaining social comparison information probably compared their exam grades to a variety of other students. Furthermore, it could be that social comparison with more successful others did not lower students' self-evaluations because they made up for a painful experience with a happy one, through the use of a downward comparison in their second choice. For these reasons, Huguet et al. (2001) included two comparison choice measures in each of the seven courses (resulting in 14 comparison choices). As expected, the participants compared slightly upward on the two choices in most courses, higher course grades were associated with higher comparison-level choices, and choosing to compare with someone who outperformed them in a course did not leave participants feeling relatively less able in that course.

Furthermore, Blanton et al. (1999) reported indirect evidence that the persons nominated as comparison targets were important in the lives of participants. Consistent with this, Huguet et al. (2001) offered direct evidence that children engaged in upward comparison with psychologically close others (at least for choice 1). According to Buunk and Ybema (1997), individuals generally avoid identification with worse-off others (with whom they try to contrast themselves) and try to identify with others doing better (and see these others as similar to themselves). Also consistent with this, and with the hypothesis that upward comparison is motivated by a desire for self-improvement (Wood, 1996), Huguet et al. (2001) found that most students reported that their performance in almost all courses might become closer to that of their more successful comparison targets in the future. Another critical issue regarding upward comparison and its impact on behaviour is whether the individuals perceive that it is possible and important to improve, these perceptions being contingent in part of feelings of control and of self-worth. Huguet et al. (2001) found that upward comparison was indeed more likely to occur when the students perceived that their degree of control over their status relative to the comparison targets was relatively high, and when the comparison dimension was important or self-relevant.

THE PRESENT STUDY

The present study expand on this prior research. Our central question is whether the tendency to compare upward can be found during the elementary school years. Some social comparison behavior is exhibited by preschool children, and interest in social comparison information increases during the early school years (e.g., Ruble & Frey, 1991). However, whether elementary school children compare upward still remains an open question. As noted by Ruble and Frey (1991), social comparison information is not used by children for abstract assessment of their abilities and behavior based on such assessment before they are able to infer abilities from overt performances, at the age of seven-eight years. If the tendency to compare upward is motivated by competence assessment and related need for self-improvement, then elementary school children (at the age of seven-eight years) should exhibit this tendency as well. As suggested by Festinger (1954), however, the upward tendency also reflects competitive pressures, and most social comparisons made by children at the age of seven-eight years still emphasize meeting developmental or age norms rather than placing high in a hierarchy (Ruble & Frey, 1991). Thus, deciding whether these children compare themselves upward with their classmates is certainly premature.

The study presented below used the same measures as those used in past research (Blanton et al., 1999; Huguet et al., 2001), but it included both elementary and middle school children (N = 339 students, overall). The participants listed on a questionnaire their two comparison-targets in various courses at the end of Trimester 2 (T2). A score of comparison-level choice was then assigned to each participant on the basis of the trimester grade their comparison-others received at the time of the nomination (T2), and the associations between participants' comparison-level choices and their own trimester grade at T2 were tested. As before, participants did not estimate their own exam grade nor did they estimate their comparison-level. Both of these scores were taken from official grade records. In addition, the participants had to answer a number of complementary questions (i.e., closeness to the comparison target and self-relevance of academic domains, taken from Huguet et al., 2001), which were adapted for younger children when needed. Because it could be difficult for elementary school children to respond in multiple courses for each comparison target, only three courses were included in the questionnaire: Mathematics, writing, and reading with children at grade levels five and six, respectively ; math, writing, and science with children at grade level seven-eight from multigrade classrooms; math, french, and science with ninth grade level children. These courses were retained because they are the most fundamental (typically viewed as such by teachers) at each grade level. Sample size with elementary school children (grade level five and six) was smaller than in past research with older (ninth grade level) children, and thus the findings from the two set of data (younger vs. older children) could be difficult to compare. This is why, we used a new sample of ninth grade level children, which was about the same size as in the three other grade levels.

Similar to previous research (Blanton et al., 1999; Huguet et al., 2001), the primary analyses were standard regression analyses conducted within each grade level. Such a strategy provided larger samples sizes for each analysis which provided a better basis for estimating relationships among the constructs of interest than conducting analyses within each individual class. Nonetheless, such a strategy does not take into account fully the nested nature of the data. In light of this shortcoming, we also conducted a series of multilevel random coefficient models in which students were treated as nested within classes. The results of these multilevel analyses are presented after the results of the more traditional regression analyses.

GRADE LEVEL FIVE

METHOD

Participants

They were 75 children at grade level five (mean age = seven years and nine months; 43 boys), who attended three French public elementary schools and four classrooms (15 to 24 pupils in each class). Only 4% of parents did not allow their children to participate.

Procedure

In the last two to three weeks of second trimester (T2), a questionnaire was administered by the regular teacher to all students in attendance on that day. The study staff met with the classroom teachers prior to collection of the data to describe how to administer the questionnaire. Students were assured of the anonymity of their responses.

Measures

Comparison-Level Choice

Participants listed two students from their classroom (hereafter called choice 1 and choice 2) with whom they typically compared their exam grades in each of the three courses. They were also told that they did not have to list anyone for courses in which they did not usually compare their exam grades. Following this, the grades of participants' comparison targets were determined. Their comparison-level choice scores in the different courses was the second trimester grade of their comparison targets in these courses because the targets were receiving this grade at the time of comparison.

Comparative Evaluation

Participants also indicated how good they were compared to most of their classmates in each course. These ratings were made using a 5-point scale (1 = much worse, 5 = much better, and 3 = the same). If participants were unsure, they could indicate this.

Complementary Measures

Closeness

In each course, participants described how much they talked to the person nominated under choice 1 using a 5-point scale (1 = much less than with my other classmates, 5 = much more than with my other classmates, 3 = as much as with my other classmates). For most French children of this age, reporting that they talk frequently to a given member of their classroom means that this member is a close friend.

Self-Relevance of the Academic Domains

Participants rated how important to themselves each academic domain was. A pre-test of a 5-point scale for this question revealed that children of this age did not understand it, therefore a 2-point scale was used (0 = not important for me, 1 = important for me). Participants were told that they could leave this item blank for courses in which they did not know how to answer.

Grades

Grades were taken from students' T2 grade official reports for each of the three courses. In the French grading system, exam grades are given on a scale ranging from 0 to 20. The typical verbal descriptors associated with grades are extremely poor from 0 to 5, poor from 6 to 9, passable to satisfying from 10 to 14, and very satisfying to excellent from 15 to 20.

RESULTS

Nomination

As in past research with older children, the vast majority of students chose to nominate comparison targets. For choice 1, rates of nomination were highest with reading (95%), and lowest with writing (91%). For choice 2, they were highest with writing (84%), and lowest with reading and math (81% in both cases). Participants who did not nominate choice 1 did not nominate choice 2 either, indicating that they understood the task. As revealed by independent t-tests, the minority of students who did not nominate comparison targets did not differ (marginally or significantly) in their exam grades from the other participants. The decision to respond was therefore unrelated to student's exam grades (which was also the case for the other three grade levels described in this paper).

Comparison Targets

Most participants, 66% for choice 1 and 54% for choice 2, chose same-sex targets (assessed across the three courses). This preference for same-sex comparisons is weaker than that reported previously (Blanton et al., 1999; Huguet et al., 2001) with ninth grade level children (around 90% and 80% for choice 1 and choice 2, respectively). It is clearly significant ($p < .001$), however, at least for choice 1. Were the comparison targets close friends, as in past research with older children? They were not. In each course, participants' closeness ratings clustered around the neutral midpoint of 3 ("I talk to my comparison target as much as with my other classmates"). No differences were found against this midpoint (assessed from one-sample t-tests). More crucial for the present paper, paired t-tests indicated that participants who nominated comparison-targets did not choose students who were doing better than they were in the different courses (see Table 1). Although the difference on exam grades between participants and their targets is compatible with a tendency to compare upward in two courses (reading and writing) for both choice 1 and choice 2, it was not significant. In math, this difference was reversed but was not significant.

Table 1. Comparison-Level Choices in the Three Courses for the Four Grade Levels

		Student		Choice 1		Choice 2	
	Course	M	SD	t	df	t	df
5th-Grade	Math	15.70	3.26	-1.60	65	-0.71	57
	Reading	15.96	2.88	1.04	50	0.32	41
	Writing	15.35	3.04	1.39	54	0.94	52
6th-Grade	Math	15.46	3.31	1.87(a)	85	0.72	84
	Reading	16.08	2.54	0.55	60	2.12*	57
	Writing	14.61	2.92	0.45	49	1.01	43
7-8th-Grade	Math	13.09	4.76	1.46	90	-0.68	80
	Science	11.55	5.10	3.70***	67	1.57	63
	Writing	08.18	7.74	1.14	91	1.52	92
9th-Grade	Math	12.58	3.40	3.23**	79	2.41*	67
	Science	12.98	3.49	2.09*	77	2.19*	65
	French	11.98	3.14	2.83**	73	2.23*	66

Note: (a)p = .06. *p <.05. **p <.01. ***p <.001 (paired t-tests, two-tailed). Positive t values indicate a tendency toward upward comparison (i.e., comparison targets' exam grade minus students' exam grade).

Association between Students' Comparison-Level Comparative Evaluations, and, Their Exam Grades

Relationships between students' comparison-level choices and their exam grades were examined with a set of three standard regression analyses (one in each course), in which students' course grades were regressed on their comparative evaluations and the grades of their comparison targets. No significant associations were found. Participants' comparative evaluations were not significantly related to their exam grades either, apart from the exception of math (see Table 2).

Table 2. Standardized Regression Coefficients for Regression of Course Grades on Students' Comparative Evaluations (CE) and Comparison-Levels (Choices 1 and 2)

	Course	CE	Choice 1	Choice 2	df	R square
5th-Grade	Math Reg Coeff	0.38**	0.13	0.12	50	.11
	Reading Reg Coeff	0.19	0.21	0.24	37	.05
	Writing Reg Coeff	0.02	0.16	0.05	45	-.03
6th-Grade	Math Reg Coeff	0.35***	0.45***	0.16	81	.31
	Reading Reg Coeff	0.10	0.20	0.42**	53	.26
	Writing Reg Coeff	0.28(a)	0.22	0.12	40	.12
7-8th-Grade	Math Reg Coeff	0.21(a)	-0.22(a)	-0.10	75	.06
	Science Reg Coeff	0.39***	0.43***	0.20*	58	.58
	Writing Reg Coeff	0.56***	0.09	0.05	86	.32
9th-Grade	Math Reg Coeff	0.55***	0.24**	0.27**	64	.66
	Science Reg Coeff	0.56***	0.18*	0.35***	62	.68
	French Reg Coeff	0.62***	0.27***	0.22**	63	.66

Significance-levels for two-tailed t-tests of regression coefficients: (a)p = .06. *p <.05. **p <.01. ***p <.001.

Association between Students' Comparison-Level Choices and Their Comparative Evaluations

Relationships between students' comparison-level choices and their perception of their relative standing were examined with another set of three standard regression analyses (one in each course), in which students' comparative evaluations were regressed on their exam grades and comparison-level choices (see Table 3). No significant relationships were found, except for choice 2 in math where a negative coefficient was significant.

Table 3. Standardized Regression Coefficients for Regression of CE on T2 Student's Course Grade and T2 Comparison-Level Choices

	Course	Student	Choice 1	Choice 2	df	R square
5th-Grade	Math					
	Reg Coeff	0.37**	0.05	-0.35**	50	.19
	Reading					
	Reg Coeff	0.20	-0.27	0.02	37	.01
	Writing					
	Reg Coeff	0.02	0.11	-0.25	45	.02
6th-Grade	Math					
	Reg Coeff	0.43***	-0.30**	-0.09	81	.14
	Reading					
	Reg Coeff	0.15	-0.06	-0.17	53	-.03
	Writing					
	Reg Coeff	0.21(a)	-0.13	0.14	40	.07
7-8th-Grade	Math					
	Reg Coeff	0.21(a)	-0.01	0.16	75	.04
	Science					
	Reg Coeff	0.60***	-0.07	0.14	58	.36
	Writing					
	Reg Coeff	0.56***	0.01	0.12	86	.32
9th-Grade	Math					
	Reg Coeff	0.81***	-0.15	-0.04	64	.50
	Science					
	Reg Coeff	0.84***	-0.06	-0.13	62	.52
	French					
	Reg Coeff	0.88***	-0.15	-0.09	63	.54

Significance-levels for two-tailed t-tests of regression coefficients: (a)p = .06. *p <.05. **p <.01. ***p <.001.

Importance Ratings

The three courses were rated as self-relevant by most participants: 81% in math, 73% in reading, and 85% in writing. Those who did not know how to answer represented 10% in math, 11% in reading, and 10% in writing. Thus, only 9% of participants in math, 16% in reading, and 5% in writing did not rate the courses as important for them personally. Excluding these participants did not change the findings. No other effects were found.

GRADE LEVEL SIX

METHOD

Participants

They were 86 children at grade level six (mean age = eight years and six month; 45 boys), who attended two public and one private elementary schools and four classrooms (17 to 28 pupils in each class). As before, only 4% of the parents did not allow their children to participate.

Procedure and Measures

The procedures and measurs for the grade level six administration were exactly the same as in grade level five.

RESULTS

Nomination

For choice 1, rates of nomination were highest with math and writing (100%), and lowest with reading (95%). For choice 2, they were highest with math (99%), and lowest with reading (89%).

Comparison Targets

Most participants, 73% for choice 1 and 70% for choice 2, chose same-sex targets (ps. < .001 across the three courses). As in grade level five, one-sample t-tests (one for each course) showed that participants' closeness ratings did not differ from the neutral midpoint (3) of the scale. In contrast with grade level five, however, paired t-tests indicated that participants who nominated comparison-targets compared or tended to compare themselves upward both in math for choice 1 and reading for choice 2 (see Table 1).

Association between Students' Comparison-Level Choices, Comparative Evaluations, and Their Exam Grades

Participants' comparison-level choices were significantly and positively related to their exam grades in math for choice 1 and reading for choice 2. Similarly, participants' comparative evaluations were significantly (or marginally significantly) and positively related to their exam grades in math and writing (see Table 2).

Association between Students' Comparison-Level Choices and Comparative Evaluations

Participants' comparison-level choice (choice1) was significantly and negatively related to their perception of their relative standing in math (see Table 3).

Importance Ratings

The three courses were again rated as self-relevant by most participants: 88% in math, 72% in reading, and 86% in writing. Those who did not know how to answer represented 5% in math, 13% in reading, and 6% in writing. Only 7% of participants in math, 15% in reading, and 8% in writing did not rate these different courses as important for them personally. Excluding these participants did not change the findings.

GRADE LEVEL SEVEN-EIGHT

METHOD

Participants

They were 98 children at grade levels seven or eight (mean age = ten years and two months; 52 boys) who attended two public elementary schools and four multigrade classrooms (with 22 to 26 pupils in each class). All parents allowed their children to participate.

Procedure and Measures

They were the same as in the first two grade levels (five and six), except the reading course, which was replaced by science. The complementary measures also differed somewhat. For closeness, participants rated their level of friendship with the students they nominated using a 5-point scale (1 = very bad, 5 = very good, and 3 = average). For self-relevance of academic domains, participants rated how important each course was using a 5-point scale (1 = very low, 5 = very high, and 3 = average). Because the classrooms were multigrades, the data for seven and eight graders were analyzed together.

RESULTS

Nomination

Nomination rates were again very high, varying from 96% for math to 100% for science and writing, for choice 1, and from 91% for math to 100% for writing, for choice 2.

Comparison Targets

As in grade level six, most participants chose same-sex targets, 80% for choice 1 and 79% for choice 2 (ps. <.001 across the three courses). The closeness ratings for this study were significantly different from the neutral midpoint (3) of the scale. In each course, participants rated choice 1 and choice 2 as close friends (ratings ranged from 3.74 for choice 2 in math to 4.45 for choice 1 in writing, all ps. < .001, assessed from one-sample t-tests). Paired t-tests showed that the closeness ratings were higher for choice 1 than for choice 2 in math and writing (ps. < .001). Another set of paired t-tests revealed that participants compared themselves significantly upward in science on choice 1 (see Table 1).

Association between Students' Comparison-Level Choices, Comparative Evaluations, and Their Exam Grades

In science, participants' comparison-level choices were significantly and positively related to their exam grades. In math, choice 1 was marginally related to students' exam grades, although the relationship was negative. Participants' comparative evaluation scores were significantly (or marginally) and positively related to their exam grades in each course (see Table 2).

Association between Students' Comparison-Level Choices and Comparative Evaluations

Student's comparison-level choices in the different courses were not significantly related to their comparative evaluations (see Table 3).

Importance Ratings

In each course, a one-sample t-test revealed that the importance ratings differed from the midpoint (3) of the scale (4.69 with math, 4.62 with science, and 4.52 with writing, all ps < .001).

GRADE LEVEL NINE

METHOD

Participants

They were 80 students (mean age = twelve years and ten months; 39 boys) in their first year of secondary school, who attended three public schools and four classrooms (with 17 to 23 students in each class). Again, all parents allowed their children to participate.

Procedure and Measures

They were the same as in grade level seven-eight, except the writing course, which was replaced by french (the same course as writing but with a different label for the older children).

RESULTS

Nominations

Nomination rates varied from 93% for science to 98% for french for choice 1, and from 81% for math to 91% for french for choice 2.

Comparison Targets

Most participants chose same-sex targets, 81% for choice 1 and 80% for choice 2 (ps. < .001 across the three courses). As in grade level seven-eight, one-sample tests based on the neutral midpoint (3) indicated that participants rated choice 1 and choice 2 as close friends (the closeness ratings ranged from 3.61 for choice 2 in math to 4.30 for choice 1 in french, all ps. < .001). Paired t-tests, however, showed that these ratings did not differ between choice 1 and choice 2. As in past research with children of the same age (Blanton et al., 1999; Huguet et al., 2001), participants compared significantly upward in each course for both choice 1 and choice 2 (see Table 1).

Association between Students' Comparison-Level Choices, Comparative Evaluations, and Their Exam Grades

Participants' comparison-level choices as well as comparative evaluations were significantly and positively related to their exam grades in each course for both choice 1 and choice 2 (see Table 2).

Association between Students' Comparison-Level Choices and Comparative Evaluations

Students' comparison-level choices in the different courses were not significantly related to their comparative evaluations (see Table 3).

Importance Ratings

Again, the importance ratings differed from the neutral midpoint (3) of the scale in each course (4.54 for french and math, and 3.71 for science, all ps < .001).

MULTILEVEL ANALYSES

The data collected in this investigation constituted a nested or hierarchical data structure because students were grouped in classes. Although the analyses reported above took this nesting into account in some ways – different grade levels were analyzed separately – these analyses do not represent best practice. From a purely technical point of view, these data should be analyzed with a series of multilevel random coefficient models (MRCM), a technique that is often and mistakenly referred to as HLM (see deLeeuw and Kreft, 1995, for a discussion of this distinction). An introduction to using MRCM to analyze data collected in groups can be found in Nezlek and Zyzniewski (1998).

Unfortunately, the present data set did not have enough classes within each grade level to provide a solid basis for a MRCM analysis (i.e., see Kreft & deLeeuw, 1999, for a discussion of sample sizes necessary for MRCM analyses). Despite this important shortcoming, for the sake of thoroughness, the four grade levels were reconceptualized as a single sample (more classes were thus available) and a series of MRCM analyses were done. In these new analyses, students were treated as nested within classes (schools were not taken into account because they were essentially confounded with classes), and through the use of dummy-coded predictors followed by tests of fixed effects (Nezlek, 2001), separate coefficients for each grade level were estimated and compared. Because the type of courses differed from one grade level to another (math, reading, and writing in the younger children; math, science, and french in the older children), only the data related to math and writing/french (taken together because they represent the same comparison dimension with a different label) could be examined.

The relationships tested by these new analyses (done using the program HLM-Version 6; Raudenbush, Bryk, Cheong, & Congdon, 2000) were similar to those tested by the previous analyses, although there were some differences primarily in terms of the number of predictors that were simultaneously included in each model. Given the same number of predictors, multilevel models estimate more parameters (error terms and their covariances) than corresponding OLS regression analyses (e.g., Nezlek, 2001).

The first set of analyses examined relationships between students' exam grades and their first comparison-level choice. For writing/french, these analyses found that the coefficient (the slope, within the terminology of multilevel modeling) between exam grades and choice 1 for ninth graders were significantly different from 0 (.40, $p < .05$), whereas none of the coefficients for the other three grade levels (.14, .13, and .01 for grade levels five, six, and seven-eight, respectively) was significantly different from 0. The pattern was similar for math. The coefficient for ninth grade level (.45) was significantly different from 0, whereas the coefficients for fifth and sixth grade levels were not (.16 and -.14, respectively). Although the coefficient for seven-eight grade level (.24) was also significantly different from 0, it was less than the coefficient for ninth grade level.

A second set of analyses examined relationships between students' grades and their second comparison-level choice. For writing/french, these analyses found that the coefficient between exam grades and choice 2 for ninth graders was marginally significantly different from 0 (.34, $p < .07$), whereas none of the coefficients for the other three grade levels (.00, .10, and .11 for grade levels five, six, and seven-eight, respectively) approached conventional levels of significance. The pattern was similar for math. The coefficient for ninth grade level (.58) was significantly different from 0, whereas none of the coefficients for the other three grade levels (-.22, -.04, and .03 for grade levels five, six, and seven-eight, respectively) approached significance.

A parallel set of analyses were done with comparative evaluation as the dependent measure instead of exam grades. The results of these analyses indicated that the relationships between comparative evaluation and comparison-level choices were stronger for the ninth grade level than for the other three grade levels. For writing/french, the coefficient between comparison evaluation and choice 1 was larger for the ninth grade level (.06) than it was for the other three grade levels (.02, .02, .00, for grade level five, six, and seven-eight, respectively), although it was not significantly different from 0 ($p = .16$). For math, the same pattern occurred. The coefficient between comparative evaluation and choice 1 was larger for the ninth grade level (.05) than it was for the other three grade levels (.00, .00, -.02 for grade level five, six, and seven-eight, respectively), although it was not significantly different from 0 ($p = .10$).

A second set of analyses examined relationships between comparative evaluation and comparison-level for choice 2. Regarding writing/french, these analyses found that the coefficient for ninth graders was not significantly different from 0 (.06), not were any of the coefficients for the other three grade levels. The pattern was somewhat different for math. The coefficient for ninth grade level (.08) was significantly different from 0, whereas the coefficients for sixth and seventh-eighth grade levels were not (.04, .01), and the coefficient for fifth grade level (-.08) was also significant but was negative.

Finally, it should be noted that the program HLM provides two sets of results, one based on robust standard errors and the other not based on robust standard errors. Robust estimates are appropriate when there is a large number of level 2 (in this case, classes) observations, and robust estimates were not appropriate for these data. This lends credence to our contention that although multilevel analyses are desirable in some ways, the small number of classes at each grade level did not provide a firm basis for a MRCM analysis. Nevertheless, although robust estimates were technically not appropriate for the present data, when they were used, the pattern of results described above was much clearer and stronger. One may think that a larger number of classes should have been sampled in order to statistically address the problem of non-independence of data. Future research would do well to consider this in the design.

Overall, however, the results of the multilevel modeling analyses are consistent with those from the standard regression analyses performed on each grade level taken separately. Whereas the relationships between students' exam grades and comparison-level choices were generally not significant in the elementary school children (with only one exception in math for choice 1 at grade level seven-eight), they were significant (choice 1) or marginally significant (choice 2) for the older children (ninth grade level). Relationships between comparison-level choices and comparative evaluations were generally not significant, even in the older (ninth grade level) children, providing further support for the idea that choosing to compare with more successful others in the different courses did not leave these children feeling relatively less able in these courses. Quite the contrary: The only exception where a significant relationship was significant in the older children, this relationship was clearly positive. The negative coefficient found in the sixth graders may be the sign that the younger children have some difficulties to benefiting from comparison with those performing well in their classroom (see the general discussion).

GENERAL DISCUSSION

Comparison-level choices have been neglected in past research on social comparison in young children. Despite this lack of attention, there is some evidence that the inferential capabilities of elementary school children are sufficiently developed to allow children to make use of comparison information for abstract assessment of their abilities and of their behavior based on such assessment (Ruble & Frey, 1991). As indicated by the present results, however, this does not mean that elementary school children compare upward, as middle school children typically do (Blanton et al., 1999; Huguet et al., 2001).

The fact that the children from grade level five did not compare themselves upward with their classmates is compatible with Ruble and Frey's (1991) earlier conclusion about the younger elementary school children: Although these children make use (as also do younger children) of social comparison, they do not yet compare as a way of determining their relative position in a hierarchy. Instead, they engage in social comparison to meet developmental or age norms. Evidence that the children from grade level five engaged in social comparison with their classmates can be found in both their low abstention rate at the nomination task and their preference for same-sex comparisons. Meisel and Blumberg (1990) reported that young children (grade level four) also show a strong preference for comparison with same-sex targets, but not with classmates nominated as close friends. This is exactly what we found here. The lack of significant upward comparisons in grade level five may have been due to the lack of statistical power. Although such a possibility cannot be ruled out, upward comparisons were still not significant in the younger children when the data of the first two grade levels were put into a single set.

Upward comparison seems a bit more obvious at grade level six, but still does not emerge in each course. This conclusion holds for the children from grade level seven-eight. Furthermore, in both grade levels, upward comparison, when it was significant, did not lower students' comparative evaluations. When they made these evaluations, participants reflected more on their own abilities than on the performances of their comparison targets. This is also true for the ninth grade level children, who compared upward in each course. As in past research with children of the same age (Blanton et al., 1999; Huguet et al., 2001), upward comparison occurred for both choice 1 and choice 2, indicating that social comparison with more successful and psychologically close others did not lower students' self-evaluations just because they made up for a potentially painful experience with a happy one (through the use of downward comparison in their second choice).

Clearly, choosing to compare with more successful others in self-relevant courses did not leave participants feeling relatively less able in these courses. Of particular interest here, Taylor and Lobel (1989) have argued that individuals are able to avoid explicit self-evaluations when they compare upward for the purpose of self-improvement. And self-improvement is indeed one of the major values of the school system, which may encourage children to compare themselves with other students in class, especially with those who perform well.

Keep in mind that students did not estimate their own exam grade nor did they estimate their comparison-level; both of these scores were taken from official grade records, and therefore the participants did not report comparing to others who had done better than them. Upward comparison was assessed here indirectly, and we believe this is a critical point. Gibbons, Blanton, Gerrard, & Buunk (2000) have shown that students were more likely to do well in school if they reported comparing their exam grades with other students who score high on tests, but they were not helped if they reported comparing to students who had done better than them (see also Nosanchuk & Erickson, 1985). Combined with these earlier findings, the presents results for ninth grade level children suggest that students can benefit from upward social comparison, but only when they do not think about the other person in a way that makes them feel worse in comparison. As suggested by Gibbons et al. (2000), one of the secrets to benefiting from comparison with those performing at a higher level is to avoid explicit reflections about one's relative merits when the comparison is being made. The negative coefficients (Table 3) sometimes found with the younger participants (choices 1 and 2 in math with children from grade levels six and five, respectively) suggest that, at this age, avoiding explicit reflections about one's relative merits may be more difficult.

Finally, Goethals and Darley (1987) noted that the school system has certain values which may lead children to compare themselves with other students in class, especially with those who perform well. Communicated implicitly in hundreds of ways on a daily basis, one central value is that good performance on the intellectual tasks set by the teacher is a good thing, specifically a better thing than less good performance. Goethals and Darley suggested that it is through the repetition of messages like this that children learn that good performances are taken by the teacher to mean high abilities and thus acquire in academic achievement settings the unidirectional drive upward postulated by Festinger (1954). The present findings provide direct evidence that this tendency indeed becomes stronger over time in the school system.

REFERENCES

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Buunk, B. P., Collins, R. L., Taylor, S. E., VanYperen, N. W., & Dakof, G. A. (1990). The affective consequences of social comparison: Either direction has it ups and down. Journal of Personality and Social Psychology, 59, 1238-1249.

Buunk, B. P., & Ybema, J. F. (1997). Social comparison and occupational stress: The identification-contrast model. In B. P. Buunk & F. X. Gibbons (Eds.), Health, coping, and well-Being: perspectives from social comparison theory (pp. 359-388). Mahwah, NJ: Lawrence Erlbaum.

Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than- others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. Journal of Personality and Social Psychology, 76, 420-430.

deLeeuw, J., & Kreft, I. G. G. (1995). Questioning multilevel models. Journal of Educational and Behavioral Statistics, 20, 171-189.

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

Gibbons, F. X., Blanton, H., Gerrard, M., & Buunk, B. (2000). Does social comparison make a difference? Optimism as a moderator of the relation between comparison-level and academic performance. Personality and Social Psychology Bulletin, 26, 637-648.

Goethals, G. R., & Darley, J. M. (1987). Social comparison theory: Self-evaluation and group life. In B. Mullen & G. R. Goethals (Eds.), Theories of group Behavior (pp. 21-47). New York: Springer-Verlag.

Huguet, P., Dumas, F., Monteil, J.-M., & Genestoux, N. (2001). Social comparison choices in the classroom: Further evidence for students' upward comparison tendency and its beneficial impact on performance. European Journal of Social Psychology, 31, 557-578.

Kreft, I. G. G., & deLeeuw, J. (1999). Introducing multilevel modeling. London: Sage.

Levine, J. M. (1983). Social comparison and education. In J. M. Levine, & M. C. Wang (Eds.), Teacher and student perceptions: Implication for learning. Hillsdale, NJ: Erlbaum.

Meisel, C. J., & Blumberg, C. J. (1990). The social comparison choices of elementary and secondary school students: The influence of gender, race and friendship. Contemporary Educational Psychology, 15, 170-182.

Nezlek, J. B. (2001). Multilevel random coefficient analyses of event and interval contingent data in social and personality psychology research. Personality and Social Psychology Bulletin, 27, 771-785.

Nezlek, J. B., & Zyzniewski, L. E. (1998). Using hierarchical linear modeling to analyze grouped data. Group Dynamics: Theory, Research, and Practice, 2, 313-320.

Nosanchuk, T. A., & Erickson, B. H. (1985). How high is up? Calibrating social comparison in the real world. Journal of Personality and Social Psychology, 48, 624-634.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. (2000). HLM. Chicago, IL: Scientific Software.

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: Lawrence Erlbaum.

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

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat : Downward evaluation and upward contacts. Psychological Review, 96, 569-575.

Wood, J. V. (1996). What is social comparison and how should we study it? Personality and Social Behavior Bulletin, 22, 520-537.

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), Self-efficacy in changing societies (pp. 202-231). Cambridge University Press.

APPENDIX

Table A. Means, Standard Deviations, and Correlation Matrix for Key Variables (Grade 5)

5th-Grade Math	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	15.92 (3.14)	--			
EC	03.44 (1.13)	.35**	--		
Choice 1	15.03 (3.25)	.17	.06	--	
Choice 2	15.49 (3.07)	.02	-.32**	.14	--
Reading	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	15.71 (3.03)	--			
EC	03.34 (1.17)	.17	--		
Choice 1	16.71 (2.72)	.16	-.20	--	
Choice 2	15.86 (2.41)	.24(a)	.08	-.06	--
Writing	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	15.30 (3.13)	--			
EC	03.47 (1.02)	.02	--		
Choice 1	16.00 (2.92)	.17	.11	--	
Choice 2	15.87 (2.97)	.05	-.26*	.06	--

Note. (a)p = .07. *p <.05. **p <.01.***p <.001.

Table B. Means, Standard Deviations, and Correlation Matrix for Key Variables (Grade 6)

6th-Grade					
Math	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	15.44 (3.32)	--			
EC	03.14 (0.62)	.28**	--		
Choice 1	16.09 (2.99)	.44***	-.13	--	
Choice 2	15.73 (2.84)	.25	-.05	.26**	--
Reading	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	16.00 (2.56)	--			
EC	03.51 (0.95)	.04	--		
Choice 1	16.13 (2.81)	.42***	-.08	--	
Choice 2	16.68 (2.69)	.52***	-.11	.54***	--
Writing	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	14.49 (2.93)	--			
EC	03.20 (0.55)	.31*	--		
Choice 1	14.93 (3.11)	.27*	-.02	--	
Choice 2	15.09 (3.48)	.27*	.22	.39**	--

Note. *p <.05. **p <.01. ***p <.001.

Table C. Means, Standard Deviations, and Correlation Matrix for Key Variables (Grade 7-8)

7-8th-Grade					
Math	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	12.97 (4.66)	--			
EC	03.33 (0.81)	.23*	--		
Choice 1	14.13 (4.67)	-.19*	.00	--	
Choice 2	12.57 (5.17)	.07	.18*	.28**	--
Science	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	11.40 (5.03)	--			
EC	02.95 (0.89)	.61***	--		
Choice 1	13.24 (5.35)	.62***	.33**	--	
Choice 2	12.43 (5.35)	.47***	.40***	.26*	--
Writing	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	08.03 (6.73)	--			
EC	02.89 (1.00)	.58***	--		
Choice 1	09.07 (6.97)	.17(a)	.12	--	
Choice 2	09.42 (6.42)	.18*	.22*	.12	--

Note. (a)p = .06. *p <.05. **p <.01. ***p <.001.

Table D. Means, Standard Deviations, and Correlation Matrix for KeyVariables (Grade 9)

9th-Grade Math	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	12.71 (3.54)	--			
EC	03.21 (0.96)	.71***	--		
Choice 1	14.04 (3.50)	.53***	.26*	--	
Choice 2	13.66 (3.69)	.60***	.37***	.54***	--
Science	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	13.00 (3.50)	--			
EC	03.14 (0.97)	.73***	--		
Choice 1	13.77 (4.08)	.45***	.27**	--	
Choice 2	13.81 (3.29)	.61***	.36**	.37***	--
French	Mean (SD)	T2 Grade	CE	Choice 1	Choice 2
T2 Grade	11.84 (3.20)	--			
EC	03.03 (0.79)	.73***	--		
Choice 1	13.17 (3.19)	.45***	.21*	--	
Choice 2	12.79 (3.42)	.44***	.24*	.27**	--

Note. *p <.05. **p <.01. ***p <.001.

AUTHORS' NOTE

We thank John Arregle (Professor in EDHEC, Nice, France) for his helpful comments on an earlier version of this article.

AUTHOR BIOGRAPHIES

Florence Dumas received her PhD in 2003 under the supervision of the second and third author. She is now associate researcher of the laboratory of cognitive psychology (Aix-Marseille I, France), and her primary research interest is social comparison and its consequences on cognitive performances. flodumas.home@wanadoo.fr.

Pascal Huguet (PhD in 1992) is director of research in the department of life sciences at the Centre National de la Recherche Scientifique (CNRS), the french equivalent of the U.S. National Science Foundation. As a full member of the laboratory of cognitive psychology (Aix-Marseille I, France), his major research interest is the social regulation of cognitive functioning. huguet@up.univ-mrs.fr.

Jean-Marc Monteil (PhD in 1978) created and directed (from 1983 to 1998) the laboratory of social psychology of cognition (associated to the CNRS at Blaise Pascal University). He is now director of higher education for the French Ministry of National Education and Research.

Claire Rastoul is postgraduate of Blaise Pascal University (Clermont-Ferrand, France). Her PhD thesis (under the supervision of the second author) is on social comparison and self-theories. rastoul@srvpsy.univ-bpclermont.fr.

John B. Nezlek (PhD in 1978) is a Professor of Psychology at the College of William and Mary in Williamsburg Virginia, USA. His major research interests are naturally occurring social interaction, daily fluctuations in psychological states, and multilevel analysis. jbnezl@wm.edu.