Classroom Factors and Student Characteristics Predicting Students' Use of Achievement Standards during Ability Self-Assessment

Douglas Mac Iver


Stable URL:
http://links.jstor.org/sici?sici=0009-3920%28198710%2958%3A5%3C1258%3ACFASCP%3E2.0.CO%3B2-0

*Child Development* is currently published by Society for Research in Child Development.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/srcd.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.
Classroom Factors and Student Characteristics
Predicting Students’ Use of Achievement Standards during Ability Self-Assessment

Douglas Mac Iver
University of Michigan

Belief in the self’s ability is a fundamental component of positive achievement motivation (e.g., Covington & Beery, 1976; Eccles & Wigfield, 1985). For this reason, researchers have been interested in the “standards of excellence” that students use in deciding whether they are competent in a given domain (e.g., Blumenfeld, Pintrich, Meece, & Wessels, 1981; Harter & Pike, 1984; Stipek & Tannatt, 1984). Veroff (1969, 1977) distinguished three broad classes of achievement standards: autonomous standards, which involve one’s own norms; social standards, which involve norms set by others; and impersonal standards, which involve the “objective” mastery criteria inherent in the task itself.

It is important to determine the effects of classroom factors on children’s use of these achievement standards. If classroom factors influence the criteria that children use to judge their ability, then developmental shifts in ability perceptions may be partly a function of systematic, grade-related changes in classroom environments (e.g., Eccles et al., 1984). In testing relations among classroom factors, student characteristics, and students’ use of achievement standards, this study will focus on two major subclasses within each standard: within autonomous standards on across-time comparisons (Masters & Keil, in press; Reuman, 1986) and across-domain comparisons (Marsh, Smith, & Barnes, 1985); within social achievement standards on social comparison (e.g., Festinger, 1954; Ruble, 1983) and use of adults’ assessments (Harter & Pike, 1984; Minton, 1979); and within impersonal standards on mastery of “everyday” tasks and of particularly difficult tasks (e.g., Mac Iver, 1986).

Although there is evidence that developmental and classroom factors produce a gradual decline in students’ ability self-concepts in elementary school (e.g., Eccles et
al., 1984), "these factors have generally not been shown to override the strong student inclination toward positive academic self-concept" (Blumenfeld, Pintrich, Meece, & Wessels, 1982, p. 416). For example, in one large sample of upper-elementary-school students, almost 83% reported that they were good at math (Mac Iver, 1986). Therefore, understanding self-assessment of math ability in early adolescence is largely a matter of understanding how students maintain a positive self-assessment. The present study examines how students use achievement standards to decide that they are "good at math."

Classroom Factors That Affect Students' Use of Achievement Standards

Different achievement standards emphasize different types of performance information. Classroom task structures, evaluation practices, and grouping patterns partly determine performance information and thus affect students' achievement standards and ability perceptions (Marshall & Weinstein, 1984; Rosenholtz & Simpson, 1984b). For example, Rosenholtz and Simpson argue that certain types of classrooms (i.e., "unidimensional," or "high resolution," classes) make it easy for students to compare their performance with classmates frequently and to detect clear performance inequalities. Students in unidimensional classes use the same materials, have little choice concerning the tasks they perform, and receive instruction as a whole class or in definite ability groups. Furthermore, the teachers in these classes assign grades frequently or otherwise provide public assessments of students' work.

Investigations comparing unidimensional and multidimensional classrooms have consistently found that classroom dimensionality affects the shared perceptions of ability differences that develop in the classroom (for a review, see Simpson & Rosenholtz, 1986). For example, students and teachers in unidimensional classrooms have higher consensus on students' relative academic rankings than do students and teachers in multidimensional classrooms. Members of unidimensional classrooms also have ability evaluations that are more dispersed and generally lower than those found in multidimensional classrooms.

While important, these investigations have not attempted to separate the effects of the various classroom organizational variables. Therefore, one cannot determine which of the classroom characteristics examined in these studies are most responsible for the observed effects. Moreover, the mediators of classroom dimensionality effects are unclear, although the effects are thought to be mediated by variations in children's reliance on different achievement standards that result from variations in the kinds of performance information they receive.

The present study fills several gaps. First, it employs multiple indicators of key classroom variables (e.g., task structure) in a large and diverse sample. Second, it tests how these different variables affect children's use of six types of achievement standards in deciding that they are good at math.

Hypothesized Effects of Task Structures and Grading Practices on Students' Use of Achievement Standards

The visibility and usefulness of social-comparative performance information is higher in classrooms in which every student works on the same task at the same time than in classrooms in which students are given several alternative assignments from which they can choose. This increased salience of social comparison is due to the tendency of undifferentiated task structures to reduce individual variation in performance across time while making inequalities in performance across students more interpretable (outcomes are more comparable), more salient, and more public (e.g., Blumenfeld et al., 1982; Marshall & Weinstein, 1984; Rosenholtz & Simpson, 1984a). In contrast, allowing students to work on their own and to make independent choices about what work to do and when to do it "reduces the amount and the interpretability of academic performance information shared among students" (Rosenholtz & Simpson, 1984b, p. 39). Thus students in classrooms with differentiated task structures may be less likely than those in other classrooms to use a social-comparative standard to decide that they are good at math (Hypothesis 1).

How might differentiated task structures influence students' reliance on adults' assessments? On the one hand, the high level of student autonomy in these structures implies that students can create their own interpretations of academic performance rather than accept the definitions of their teachers (Rosenholtz & Simpson, 1984a). On the other hand, because differentiated task structures make it more difficult for students to evaluate their own performance through social comparison, students may compensate by relying more heavily on adults' evaluations, but only if these evaluations are readily available. Grades summarize teacher assessments in a compact and compelling way. When grades
are assigned frequently in classrooms that have differentiated task structures, the grades may override the multiple performance dimensions created by the task structures (Simpson & Rosenholtz, 1986) and make adult assessments quite salient. In contrast, if grades are assigned infrequently in this type of classroom, students' use of adults' assessments may be quite low. That is, in the absence of grades or other formal evaluations, students may find it difficult to judge how their teacher rates their performance in differentiated, multitask environments. In these environments, teachers' informal evaluations are less comparable and less visible than in other organizational arrangements (Bossert, 1979; Marshall & Weinstein, 1984). Thus students' use of adults' assessments is hypothesized to be a joint function of both task structure and frequency/emphasis of grading; use of this achievement standard may be particularly low when task structures are differentiated and grades are infrequent but may be particularly high when differentiated structures and frequent grades coexist (Hypothesis 2).

The policy of assigning grades forces students to attend to information that is inherently stratifying (Rosenholtz & Simpson, 1984a). Although performance feedback of some kind may be present in all classrooms, in classrooms in which grades are given on each assignment, one's level of task mastery and one's standing in math compared to classmates are particularly unmistakable. Thus it will be easier for low-ability students to believe that they outperform others and master their math assignments when no grades are given. Conversely, frequent grades should reinforce the conclusion of highly able students that they are mastering assignments and outperforming others. Thus low-ability students should show greater reliance on mastery of ordinary tasks and on social comparison for deciding that they are good at math when grades are infrequent, and the opposite pattern should characterize high-ability students (Hypothesis 3).

**Hypothesized Effects of Ability Grouping on Students' Use of Achievement Standards**

There are two major types of ability grouping, namely, within-classroom grouping, where students of different ability levels are in one classroom but are divided into smaller groups on the basis of ability, and between-classroom grouping, where students are assigned to separate classrooms on the basis of ability. Between-classroom ability grouping in mathematics is uncommon in elementary school but is the most common instructional arrangement at the beginning of junior high school (Eccles et al., 1984). In both types of grouping, a student's group assignment seldom changes across time (Hallinan & Sorensen, 1983; Schafer & Olexa, 1971).

Ability grouping affects a wide range of student outcomes. For example, within-classroom grouping increases the variance in the achievement distribution of a class (Hallinan & Sorensen, 1983) and influences the formation of friendships (Bossert, 1979; Hallinan & Sorensen, 1985). Furthermore, ability-group placement serves as an important basis for social comparison. For example, children in within-classroom ability groups rely heavily on group placement as a basis for judging the relative reading abilities of their classmates (Eder, 1983; Filby & Barnett, 1982). Students assigned to a high-ability group may take their assignment as evidence of their superiority to their classmates and so may be particularly likely to base their self-perceptions of competence in math on a social-comparison achievement standard. Conversely, students assigned to a low-ability group may be particularly unlikely to decide on the basis of social comparison that they are good at math (Hypothesis 4). These effects of group placement may be stronger in within-classroom than in between-classroom ability groups because, in within-classroom groups, students have daily exposure to the other groups (Hypothesis 5).

Ability grouping is also expected to affect children's use of task mastery information during self-assessment. One common rationale for ability grouping is that it allows teachers to tailor the difficulty of assignments to the ability level of students (e.g., Sorensen & Hallinan, 1986). In classrooms in which neither ability-grouping nor individualized instruction are used, high-ability students are likely to receive a high proportion of tasks that are easy for them to master. Ability grouping makes it more feasible for teachers to give these students challenging assignments. One side effect of giving these students more challenging tasks may be to depress their use of success on daily assign-

---

1 Even mastery-based grades encourage students to compare their performances and to develop a consensus concerning how students rank in terms of ability.
ments as a basis for deciding that they are good at math because they will be succeeding less impressively on the assignments they receive (Hypothesis 6). The opposite effect may be found for low-ability children if they encounter easier assignments in grouped than in ungrouped, nonindividualized classrooms.

**Student Characteristics That Affect Students' Use of Achievement Standards**

Students' own characteristics are likely to be associated with self-assessment processes. For example, sex differences in self-assessment strategies may account in part for females having significantly lower ability perceptions in mathematics than males have beginning in early adolescence even though they actually perform as well as, if not better than, males (Marsh et al., 1985; Meece, Eccles, Kaczala, Goff, & Futterman, 1982). Furthermore, the fact that even low-performing students tend to perceive themselves as basically good at math (e.g., Mac Iver, 1986; Reuman, 1986) may indicate that low-performing students tend to adapt their self-assessment strategies to emphasize information indicating that they are able. The present study explores how student sex and student talent and performance (as rated by teachers) are related to students' reliance on different achievement standards in reaching the conclusion that they are good at math.

**Method**

**Sample**

The present study is based on data collected in fall 1983 as part of the Transitions in Early Adolescence project. Twelve school districts with varying educational practices were recruited for the Transitions project. The school districts are located near a major metropolitan area in the Midwest and serve middle-income communities. Almost 90% of the students in these districts are Caucasian. All teachers in these districts who taught fifth or sixth graders who were scheduled to make the transition to middle/junior high school the following year were invited to participate. Ninety-five percent of these teachers accepted. In this way, 143 classrooms were recruited for the study: 14 fifth-grade classrooms, 107 sixth-grade classrooms, and 22 classrooms containing students of more than one grade level. Students participated on a voluntary basis; 79% (3,248/4,110) of all enrolled students participated.

**Case Selection**

Children who answered “yes” to the question, “Are you good at math?” were the only ones asked to fill out the measures that served as dependent variables in this investigation. Therefore, the analyses reported here are necessarily limited to those 2,704 students (out of 3,204 respondents) who reported that they were good at math. Consistent with past research on classroom dimensionality effects (e.g., Rosenholtz & Wilson, 1980), only students in “single-grade-level” classrooms were analyzed; this restriction eliminated another 380 students. Because a representative sample of students from a classroom is required to gain an accurate picture of achievement standards in that classroom, classes with less than a 70% student participation rate were also excluded; this criterion eliminated 412 additional students. Of the 1,912 remaining students, 1,570 (80%) had complete data on all independent and dependent measures.

**Measures**

**Dependent Variables: Achievement Standards Used to Assess Ability by Children Who Think They Are Good at Math**

Students were first asked to respond “yes” or “no” to the question “Are you good at math?” Students who answered “yes” were then asked to “tell us which of the following things have helped you decide that you’re good at math” by rating items measuring their reliance on Adults’ Assessments, Social Comparisons, Across-Domain Comparisons, Across-Time Comparisons, Mastery of Ordinary Tasks, and Mastery of Difficult Tasks (see sample items in Appendix 1). Confirmatory factor analyses suggest that each item loads solely and substantially on the factor it is supposed to measure and that students distinguish among these six positively correlated factors (for more details, see Mac Iver, 1986). The dependent variables in this study were created by summing the three items defining each factor.

**Classroom Characteristics**

**Task structure and deemphasis of grades.**—Items measuring how often the teachers in this sample used differentiated, autonomy-promoting task structures and how often they assigned and emphasized grades were obtained from the classroom environment inventory (Midgley & Feldilhafer, 1984) that was filled out by teachers (see Appendix 2). Confirmatory factor analyses suggest that each item measures its intended factor (Mac Iver, 1986). The maximum-likelihood estimate of the true correlation between the Task Structure and Emphasis on Grades factors is \(-.11 (z = -1.0)\).
Two composite variables were created by summing the component items defining each factor. High scores on the Task Structure composite signify undifferentiated, autonomy-inhibiting task structures. High scores on the Deemphasis of Grades composite signify classrooms in which grades are infrequent and not emphasized by the teacher.

Ability grouping.—Ability-grouping practices were measured through teacher reports. None of the teachers in the subsample considered in this paper reported using individualized instruction methods. However, ability grouping was used in 35 classrooms. Teachers in these classrooms were asked to report the ability level and type of each student’s group. On the basis of this information it was possible to classify students into one of the following seven categories: ungrouped; below-average, average, or above-average between-classroom grouped; or below-average, average, or above-average within-classroom grouped.

Student Characteristics: Sex, Performance Relative to Classmates, and Talent

Student sex was measured through self-report. Student performance relative to classmates and talent in math were measured through teacher report using items developed by Parsons (1980): “Compared to other students in this class, how well is this student performing in math?” (1 = “near the bottom of the class,” 5 = “near the top of the class”) and “How much natural mathematical talent does this student have?” (1 = “very little math talent,” 7 = “a lot of math talent”).

Results

Overview of Analysis Strategy

The goal of the analyses was to estimate effects of three classroom factors (task structure, emphasis on grades, and ability grouping) and three student characteristics (sex, math talent, and math performance) on students’ use of each of the six major types of achievement standards in deciding that they were good at math. This goal was reached by estimating six different series of multiple-regression models, one series for each dependent variable. The first model estimated in each series contained all six possible main effects and all 15 possible two-way interactions. Each interaction in this model was tested for significance. Next, a second model that omitted all nonsignificant interactions was estimated. Following usual standards (e.g., Pedhazur, 1982, p. 440), a .10 level of significance was used in testing interactions. This led to the initial retention of 13 interactions in the six analysis sequences.

Although the number of interactions found here is only 1.4 times larger than the number that would be expected by chance, it should be noted that ordinary multiple regression, when applied to nonexperimental data, typically underestimates the prevalence and size of interactions (e.g., Morris, Sherman, & Mansfield, 1986). Two precautions were taken to guard against the interpretation of “chance” interactions. First, three of the 13 interactions initially retained were eliminated from the final regression models because the interactions failed to remain significant once the first set of nonsignificant interactions had been removed. Second, only interactions that appeared consistently in more than one analysis sequence are discussed in the discussion section of this paper.

After nonsignificant interactions were eliminated, the significance of each main effect was tested (using $\alpha = .05$). Nonsignificant main effects were deleted from the model unless they were nested in a significant interaction. Then the model was reestimated. The final regression models from each of the six analysis sequences are summarized in Table 1. Whenever multiple comparisons were needed to interpret effects in these models, Bonferroni tests were used. That is, within each family of comparisons considered, the error rate was controlled by adjusting the critical alpha level to take account of the number of comparisons in the family.

Although the hypotheses in this investigation were stated in language that befitted causal predictions, causality cannot be inferred from this cross-sectional, nonexperimental data set. Instead, the analyses reported here are designed to test whether the
TABLE 1

EFFECTS OF SEX (S), ABILITY (A), EMPHASIS (E), MATH TALENT (T), UNDIFFERENTIATION OF TASK STRUCTURES (U), AND MATH PERFORMANCE (P) IN FINAL REGRESSION MODELS

<table>
<thead>
<tr>
<th>Effect</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analysis 1: Social Comparison (R² = .050</strong>*)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S × A</td>
<td>6,1552</td>
<td>2.43</td>
<td>.02</td>
</tr>
<tr>
<td>E × T</td>
<td>1,1552</td>
<td>3.47</td>
<td>.06</td>
</tr>
<tr>
<td>S</td>
<td>1,1552</td>
<td>.74</td>
<td>.39</td>
</tr>
<tr>
<td>A</td>
<td>6,1552</td>
<td>1.08</td>
<td>.37</td>
</tr>
<tr>
<td>E</td>
<td>1,1552</td>
<td>2.88</td>
<td>.09</td>
</tr>
<tr>
<td>T</td>
<td>1,1552</td>
<td>3.76</td>
<td>.05</td>
</tr>
<tr>
<td>P</td>
<td>1,1552</td>
<td>13.78</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Analysis 2: Across-Time Comparison (R² = .034</strong>*)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S × A</td>
<td>6,1553</td>
<td>3.53</td>
<td>.002</td>
</tr>
<tr>
<td>T × P</td>
<td>1,1553</td>
<td>8.40</td>
<td>.004</td>
</tr>
<tr>
<td>S</td>
<td>1,1553</td>
<td>.72</td>
<td>.40</td>
</tr>
<tr>
<td>A</td>
<td>6,1553</td>
<td>1.92</td>
<td>.07</td>
</tr>
<tr>
<td>T</td>
<td>1,1553</td>
<td>7.86</td>
<td>.005</td>
</tr>
<tr>
<td>P</td>
<td>1,1553</td>
<td>10.31</td>
<td>.001</td>
</tr>
<tr>
<td><strong>Analysis 3: Across-Domain Comparison (R² = .021</strong>*)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E × A</td>
<td>6,1554</td>
<td>1.91</td>
<td>.08</td>
</tr>
<tr>
<td>E × U</td>
<td>1,1554</td>
<td>9.13</td>
<td>.003</td>
</tr>
<tr>
<td>E</td>
<td>1,1554</td>
<td>9.36</td>
<td>.002</td>
</tr>
<tr>
<td>A</td>
<td>6,1554</td>
<td>2.31</td>
<td>.03</td>
</tr>
<tr>
<td>U</td>
<td>1,1554</td>
<td>7.40</td>
<td>.007</td>
</tr>
<tr>
<td><strong>Analysis 4: Adults’ Assessments (R² = .035</strong>*)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S × A</td>
<td>6,1552</td>
<td>2.40</td>
<td>.03</td>
</tr>
<tr>
<td>E × U</td>
<td>1,1552</td>
<td>5.07</td>
<td>.02</td>
</tr>
<tr>
<td>S</td>
<td>1,1552</td>
<td>.75</td>
<td>.38</td>
</tr>
<tr>
<td>A</td>
<td>6,1552</td>
<td>2.61</td>
<td>.02</td>
</tr>
<tr>
<td>E</td>
<td>1,1552</td>
<td>6.03</td>
<td>.01</td>
</tr>
<tr>
<td>U</td>
<td>1,1552</td>
<td>5.35</td>
<td>.02</td>
</tr>
<tr>
<td>P</td>
<td>1,1552</td>
<td>14.80</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Analysis 5: Mastery of Ordinary Tasks (R² = .069</strong>*)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E × T</td>
<td>1,1559</td>
<td>4.16</td>
<td>.04</td>
</tr>
<tr>
<td>E</td>
<td>1,1559</td>
<td>2.73</td>
<td>.10</td>
</tr>
<tr>
<td>T</td>
<td>1,1559</td>
<td>7.35</td>
<td>.007</td>
</tr>
<tr>
<td>A</td>
<td>6,1559</td>
<td>2.85</td>
<td>.009</td>
</tr>
<tr>
<td>P</td>
<td>1,1559</td>
<td>20.33</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Analysis 6: Mastery of Difficult Tasks (R² = .044</strong>*)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S × A</td>
<td>6,1554</td>
<td>2.04</td>
<td>.06</td>
</tr>
<tr>
<td>S</td>
<td>1,1554</td>
<td>2.68</td>
<td>.10</td>
</tr>
<tr>
<td>A</td>
<td>6,1554</td>
<td>.62</td>
<td>.71</td>
</tr>
<tr>
<td>T</td>
<td>1,1554</td>
<td>5.54</td>
<td>.02</td>
</tr>
<tr>
<td>P</td>
<td>1,1554</td>
<td>5.05</td>
<td>.02</td>
</tr>
</tbody>
</table>

**NOTE.—Nonsignificant main effects are listed if they are nested in a significant interaction.**

**p < .01.

***p < .001.
hypotheses are consistent or inconsistent with the relations actually observed in the data.

**Analysis 1: Social Comparison**

Two interactions were retained in the final model, namely, a Sex × Ability Group interaction and a Deemphasis of Grades × Talent of Student interaction.

To interpret the Sex × Ability Group interaction, adjusted means on the dependent variable for each sex in each type of ability group were obtained after controlling for the main effects and interactions in the final model that do not involve sex or group. Two families of Bonferroni comparisons among these means were made: (1) the simple main effects of sex for each level and type of ability grouping and (2) the simple main effects of type of ability grouping for each level of ability group and sex of child. The only significant sex difference was found in above-average within-classroom ability groups: boys in these groups relied on social comparison more than did girls; boys’ M = 9.99, girls’ M = 8.80, F(1,1552) = 9.75, p < .05. The only effect of ability-group type that approached significance indicated that boys in above-average between-classroom groups (M = 9.12) tended to rely on social comparison less than did boys in above-average within-classroom groups, F(1,1552) = 6.34, p < .10. The results of both families of comparisons emphasize the relatively higher use of social comparison by boys assigned to above-average within-classroom groups.

In Hypothesis 3, it was hypothesized that low-ability students would show greater reliance on social comparison as a basis for deciding they are good at math when grades are infrequent than when they are frequent. The opposite pattern was predicted for high-ability students. The Deemphasis of Grades × Talent of Student interaction is consistent with this hypothesis (see Fig. 1). In studying this figure, consider first those students whom the teacher rates as having “very little math talent.” When grades are de-emphasized and infrequent, the reliance on social comparison displayed by these students deviates in a positive direction from the amount that would be predicted on the basis of their relative performance, ability group, or other variables in the final regression model. However, when grades are frequently assigned and emphasized, these students show less reliance on social comparison than would otherwise be predicted. In sum, for students who are below the midpoint in talent, use of social comparison as a basis for deciding that they are good at math increases as emphasis on grades decreases. In contrast, for students above the midpoint in talent, use of social comparison increases as grades become more frequently assigned and emphasized.³

There was a main effect of students’ relative performance on use of a social-comparative achievement standard. Students who are rated as performing well relative to their classmates are more likely than others to indicate that they have decided on the basis of social comparison that they are good at math (b = .28, β = .28).

**Analysis 2: Across-Time Comparison/Improvement**

The final regression model for this dependent variable contains the Sex × Ability Group and the Talent of Student × Performance of Student interactions as well as the four main effects nested in these interactions (see Table 1).

Comparisons among the adjusted means involved in the Sex × Ability Group interaction indicate that boys in an above-average within-classroom group (M = 10.43) rely on improvement more than do girls in that same group (M = 9.46) and also more than do boys in an above-average between-classroom group, M = 9.58, F’s(1,1553) > 8.80, p’s < .05.

Figure 2 displays the interactive effects of talent and performance (as measured through teacher reports) on use of improvement after controlling for the interactive effects of sex and ability grouping on this dependent variable. For all adolescents except the most talented, use of improvement as a basis for deciding that one is good at math increases at higher performance levels. For the most talented children (rated 6 or 7 on the seven-point scale), use of improvement remains the same or decreases slightly at higher performance levels.

**Analysis 3: Across-Domain Comparison**

The analyses (see Table 1) indicate that the more a teacher de-emphasizes grades the less his or her students use across-domain comparisons in inferring math ability (b = −.43, β = −.62). Undifferentiated task struc-
Fig. 1.—Interactive effects of de-emphasis of grades and talent of student on use of social comparison.

Fig. 2.—Interactive effects of talent of student and performance of student on use of improvement.
tures are also associated with less use of these comparisons \((b = -0.23, \beta = -0.28)\). However, these main effects are qualified by an interaction. The use of across-domain comparisons is highest in classes with differentiated task structures—if grades are frequent—but is lowest in such classes if grades are infrequent.

The analyses also indicate a main effect of ability grouping that is qualified by a Deemphasis of Grades \(\times\) Ability Grouping interaction. Children in below-average between-classroom groups use across-domain comparisons more than do children in above-average between-classroom groups, and this group difference is larger when grades are emphasized than when they are deemphasized.

**Analysis 4: Adults’ Assessments**

The only main effect in the final model not nested in an interaction is the effect of students’ performance in math relative to their classmates (see Table 1). Students who are performing well relative to their classmates are more likely to indicate they have decided that they are good at math on the basis of adults’ assessments of their ability than are students who are performing poorly \((b = .17, \beta = .10)\).

The Sex \(\times\) Ability Grouping interaction is similar to those found in Analyses 1 and 2. Boys in above-average within-classroom groups used adult assessments more than did girls in those groups \((M’s = 10.11 \text{ and } 9.19)\) and more than did boys assigned to between-classroom groups of the same level, \(M = 9.20, F’s(1,1552) > 8.70, p’s < .05\).

The Deemphasis of Grades \(\times\) Undifferentiation of Task Structure interaction is consistent with Hypothesis 2 (see Fig. 3). Each line in Figure 3 represents the relation between undifferentiation of task structures and usage of adults’ assessments at a particular level of emphasis on grades. The use of adults’ assessments is lowest when grades are deemphasized in highly differentiated task structures and is highest when emphasized grades and differentiated task structures coexist.

**Analysis 5: Mastery of Ordinary Tasks**

The final regression model contains two main effects that are not nested in interactions, that of performance and that of ability grouping (see Table 1). The performance effect reflects the positive relation between math performance as assessed by teacher and use of success on ordinary tasks as a basis for inferring math ability \((b = .26, \beta = .18)\).
The adjusted means used to assess ability-grouping effects were ungrouped = 10.15, between-classroom below average = 10.31, between-classroom average = 10.43, between-classroom above average = 9.79, within-classroom below average = 9.87, within-classroom average = 10.22, and within-classroom above average = 10.06. Two families of Bonferroni comparisons among these means were constructed. The first family of comparisons included all six possible comparisons of ungrouped students with grouped students. Students in between-classroom above-average groups used success on ordinary tasks less heavily than did ungrouped students, $F(1,1559) = 9.04, p < .05$. None of the other comparisons in this family was significant. The second family of comparisons involved comparison of the within-classroom groups with each other and of the between-classroom groups with each other. In between-classroom groups, the adjusted mean of the above-average group was significantly lower than the adjusted mean of the average group, $F(1,1559) = 12.28, p < .05$, and marginally lower than the adjusted mean of the below-average group, $F(1,1559) = 6.34, p < .10$. There were no significant differences among the adjusted means of the within-classroom groups.

The Deemphasis of Grades × Math Talent interaction is consistent with Hypothesis 3; infrequent emphasis on grades increases the use of success on ordinary tasks by untalented students but decreases the use of this standard by talented students. Although students’ talent levels are generally positively related to their use of success on ordinary tasks, this relation is much stronger when grades are salient than when they are not.

**Analysis 6: Mastery of Difficult Tasks**

Significant main effects indicate that talented students use mastery of difficult tasks as a basis for their math-ability perceptions more often than do untalented students ($b = .14, \beta = .10$) and that students who are performing well relative to their classmates use the same standard more than do students who are performing poorly ($b = .16, \beta = .09$; see Table 1). The Sex × Ability Grouping interaction is similar in nature to those found in Analyses 1, 2, and 4.

**Discussion**

Past research on classroom-dimensionality effects has assumed that the infrequent grades and differentiated task structures in multidimensional classrooms decrease children’s reliance on social comparison by making global inequalities in performance among students less salient and less visible than in other classrooms. However, in this study the participants did not report less overall reliance on social comparison during self-assessment in classrooms having infrequent grading and differentiated tasks. Instead, the results suggest that the combination of infrequent grades and differentiated task structures decreases children’s reliance on adults’ assessments and across-domain comparisons. That is, in Analyses 3 and 4, an inspection of the Deemphasis of Grades × Undifferentiation of Task Structure interactions revealed that the lowest use of adults’ assessments and across-domain comparisons occurred when infrequent grades and differentiated task structures were combined.

These results from Analyses 3 and 4 fit nicely with the literature on classroom dimensionality. In regard to adults’ assessments, Simpson and Rosenholtz (1986) have consistently found that adults’ assessments have less of an effect on children’s self-assessments when classrooms are multidimensional. As argued earlier, students may find it difficult to judge how their teacher rates their performance in these types of classrooms. In regard to across-domain comparisons, Rosenholtz and Simpson (1984a) would expect multidimensional mathematics classrooms to prompt students to develop a differentiated conception of their math ability rather than a conception that is generalized and singular. Without a singular self-concept of math ability, across-domain comparisons may be difficult.

It should be emphasized that the decreased use of adults’ assessments and across-domain comparisons discussed in the preceding paragraphs occurred only when task structures were differentiated in classrooms where grades were infrequent. Students displayed particularly heavy reliance on these two achievement standards in classrooms in which these differentiated task structures were accompanied by frequent grades. Grades are comparable criteria that communicate teacher evaluations in a fairly standardized way across a variety of subject areas. Grades apparently can override the multiple performance dimensions created by differentiated task structures and can focus students on adults’ assessments and on how these assessments compare across subject areas.

As noted earlier, Rosenholtz and Simpson (1984b) claim that the policy of assigning grades forces students to “attend to information structured to be inherently stratifying”
the higher students performed relative to their classmates, the more they reported relying on social comparison, adults' assessments, and mastery of ordinary and difficult tasks in deciding that they were good at math. This may indicate that students who are performing better than others in math are aware that their inference that they are good at math is strongly supported by these achievement standards, whereas children who are performing less well at math are aware that their judgment that they are good at math is only moderately supported by these standards. The only exceptions to this general pattern occurred for use of across-time comparisons and across-domain comparisons. Highly performing students, if they have high levels of natural math talent, do not usually point to improvement as a basis for their ability inferences. Perhaps they have been performing so well for so long that their consistent improvement is not very salient to them as a basis for judging their ability. Furthermore, highly performing students are not more likely than are others to use across-domain comparisons; one's performance relative to others is largely irrelevant to the use of internal, across-domain comparisons. Instead, one's performance in math relative to one's performance in other subjects is probably the crucial predictor of use of across-domain comparisons.

Although there were significant effects involving ability grouping in every analysis, they were, without exception, more complicated than hypothesized. The most prevalent finding was a Sex × Ability Group interaction. Boys assigned to an above-average within-classroom group display an elevated reliance on social comparison, adults' assessments, mastery of difficult tasks, and improvement as bases for their ability judgments. In contrast, girls assigned to this type of group do not display above-average reliance on these achievement standards. This suggests that there may be a consistent sex difference in how upper-elementary-school students react to being assigned to an above-average within-classroom group. Because this sex difference is discussed in detail elsewhere (MacIver, 1986), it will not be pursued further here.

It was predicted that, as a result of receiving more challenging tasks, students placed in above-average ability groups would show lower usage of success on daily tasks than would ungrouped students of similar ability. This hypothesis was supported in between-classroom groups but not in within-classroom groups. Under within-classroom grouping, "high-group" students can directly observe the easier curriculum and lower mastery levels of other groups. This may help them to maintain a belief that their own mastery level is high even as they struggle with daily assignments that are more difficult than those used in the typical ungrouped classroom.

The most common effect of students' math performance on use of achievement standards was a simple positive main effect; the higher students performed relative to their classmates, the more they reported relying on social comparison, adults' assessments, and mastery of ordinary and difficult tasks in deciding that they were good at math. This may indicate that students who are performing better than others in math are aware that their inference that they are good at math is strongly supported by these achievement standards, whereas children who are performing less well at math are aware that their judgment that they are good at math is only moderately supported by these standards. The only exceptions to this general pattern occurred for use of across-time comparisons and across-domain comparisons. Highly performing students, if they have high levels of natural math talent, do not usually point to improvement as a basis for their ability inferences. Perhaps they have been performing so well for so long that their consistent improvement is not very salient to them as a basis for judging their ability. Furthermore, highly performing students are not more likely than are others to use across-domain comparisons; one's performance relative to others is largely irrelevant to the use of internal, across-domain comparisons. Instead, one's performance in math relative to one's performance in other subjects is probably the crucial predictor of use of across-domain comparisons.

There are several limitations of the present study that should be acknowledged. First, the study does not address the effect of task structure on children's use of achievement standards in those extremely rare classrooms in which tasks are often or always differentiated and autonomy-promoting. The highest level of differentiation observed in this sample was a 10 on the 20-point scale (i.e., halfway between "often" and "sometimes"). The failure to find a main effect of task structure on students' use of social comparison may indicate that, even in classrooms in which tasks are sometimes differentiated, students still encounter enough "unidimensional" situations to allow them to make interpretable social comparisons.

Second, this study does not consider upper-elementary-school students who have concluded that they are incompetent in mathematics. If the effects of classroom organization on self-assessment processes is to be fully understood, then students' use of achievement standards in deciding that they are incompetent in math must also be investigated.

Third, classroom organization influences teacher perceptions of students' capabilities (Rosenholtz & Simpson, 1984b). Due to this "shared variance" between teacher ratings of student talent and classroom factors, the testing of classroom organizational effects while
controlling for the effect of teacher ratings of student talent may lead to an underestimation of the importance of classroom organization.

Finally, the number of subjects in some of the specific ability-grouping conditions was rather small. Thus some tests of the effects of ability-group type and level on use of achievement standards had less than desirable statistical power.

Environmental psychology is just beginning to have a significant effect on developmental psychology (e.g., Higgins & Parsons, 1983). Investigations, such as the present study, that explore the effect of classroom environments on children are helpful in identifying situational factors that may be responsible for developmental changes. Although the present study was cross sectional, it explored the effect of classroom environment features that are known to change systematically across the school years. As children make the transition into junior high school, they encounter greater use of undifferentiated task structures and between-classroom ability grouping and experience an increasing emphasis on grades (Eccles et al., 1984; Gronlund, 1974). These systematic changes in classroom environments may be partially responsible for the developmental changes that are observed in children’s achievement-related beliefs, attitudes, and behaviors in junior high school. For example, although a large proportion of sixth-grade students consider themselves to be good at math, students’ perceptions of their ability in mathematics declines in junior and senior high school (Eccles et al., 1984). The results of the present study help us to understand the factors that will eventually lead many of the less talented students to conclude that they are poor at math. Specifically, as emphasis on math grades increases in junior and senior high school, it will become increasingly difficult for less talented students to conclude that they are good at math. Although at first these less talented students may be able to use substitute criteria such as across-domain comparisons to maintain the illusion that they are good at math, it is likely that these substitute criteria will eventually be rejected as not credible. The conception of ability that is institutionalized in our classrooms and society emphasizes social-comparative criteria in determining who is able (Rosenholtz & Simpson, 1984b). Thus, as children mature, they develop “a realization that social comparison is necessary for adequate ability inferences” (Nicholls, 1983, p. 192). Although mature self-assessors also use impersonal mastery-oriented standards, the results of this study indicate that a high emphasis on grades inhibits less talented students’ use of simple mastery just as it undercuts their use of social comparison. Thus, as grades become increasingly emphasized across the school years, less talented students will find it increasingly difficult to escape the conclusion that they are poor at math.

The results of this study emphasize the importance of distinguishing among different classroom environment features, types of achievement standards, and categories of students in investigating the influence of classroom factors on students’ ability self-assessments. For example, the effects of classroom organization on students’ use of achievement standards frequently depend on the specific classroom organizational variable, achievement standard, and type of student under consideration. An understanding of self-assessment may require a complex interactional model of the type advocated by Marshall and Weinstein (1984). The present study provides some of the critical data needed to make construction of such a model possible.

Appendix 1

Sample Items from Factors Measuring Achievement Standards That Children Have Used to Decide That They Are Good at Math

Factor 1: Adults’ Assessments ($\alpha = .70$)
Knowing what teacher thinks about my ability . . .
Factor 2: Social Comparison ($\alpha = .70$)
Knowing how well I do compared to other kids my age . . .
Factor 3: Mastery of Ordinary Tasks ($\alpha = .65$)
Knowing I usually understand my assignments . . .
Factor 4: Mastery of Difficult Tasks ($\alpha = .76$)
Knowing I understand even the hardest types of math . . .

4 Students indicated whether each standard “has helped me decide I’m good at math” on a four-point scale ranging from “not at all true” (1) to “very true” (4). Each factor was measured by three items. For further information, see Mac Iver (1986).
1270  Child Development

Factor 5: Across-Domain Comparisons
\( (\alpha = .80) \)
Knowing how well I do in math compared to other subjects . . .

Factor 6: Across-Time Comparisons/Improvement \( (\alpha = .67) \)
Knowing I understand math better each year . . .

Appendix 2

Composition of Factors Measuring Classroom Characteristics\(^5\)

Factor 1: Task Structure \( (\alpha = .79) \)
Most students in this class use the same math textbooks and materials. (Reverse scoring was used for this item.)

Students are given several alternative math assignments from which they can choose the ones to work on for that period.

Students are given the opportunity to work on their own for several days before checking with me.

Students work on a variety of different math activities and assignments at the same time in this class.

Factor 2: Deemphasis of Grades \( (\alpha = .70) \)
I give grades on math homework assignments.

I give grades on math classwork.

I stress the importance of getting good grades in math.

Students are asked to show low grades or unsatisfactory work to their parents.

References


\(^5\) Items are on a response scale ranging from 1 ("always") to 5 ("never").


Classroom Factors and Student Characteristics Predicting Students' Use of Achievement Standards during Ability Self-Assessment
Douglas Mac Iver
Stable URL: http://links.jstor.org/sici?sici=0009-3920%28198710%2958%3A5%3C1258%3ACFASCP%3E2.0.CO%3B2-0

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

The Formation and Role of Self Perceptions of Ability in Elementary Classrooms
Phyllis C. Blumenfeld; Paul R. Pintrich; Judith Meece; Kathleen Wessels
The Elementary School Journal, Vol. 82, No. 5, Special Issue: Students in Classrooms. (May, 1982), pp. 400-420.
Stable URL: http://links.jstor.org/sici?sici=0013-5984%28198205%2982%3A5%3C400%3ATFAROS%3E2.0.CO%3B2-9

Ability Grouping and Students' Academic Self-Concepts: A Case Study
Donna Eder
Stable URL: http://links.jstor.org/sici?sici=0013-5984%28198311%2984%3A2%3C149%3AAGASCS%3E2.0.CO%3B2-R

Student Perceptions of "Better Readers" in Elementary Classrooms
Nikola N. Filby; Bruce G. Barnett
The Elementary School Journal, Vol. 82, No. 5, Special Issue: Students in Classrooms. (May, 1982), pp. 435-449.
Stable URL: http://links.jstor.org/sici?sici=0013-5984%28198205%2982%3A5%3C435%3ASPO%22RI%3E2.0.CO%3B2-Q
The Formation and Stability of Instructional Groups
Maureen T. Hallinan; Aage B. Sørensen
Stable URL: http://links.jstor.org/sici?sici=0003-1224%28198312%2948%3A6%3C838%3ATFASOI%3E2.0.CO%3B2-Q

Ability Grouping and Student Friendships
Maureen T. Hallinan; Aage B. Sørensen
Stable URL: http://links.jstor.org/sici?sici=0002-8312%28198524%2922%3A4%3C485%3AAGASF%3E2.0.CO%3B2-3

The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children
Susan Harter; Robin Pike
Stable URL: http://links.jstor.org/sici?sici=0009-3920%28198412%2955%3A6%3C1969%3ATPSOPC%3E2.0.CO%3B2-6

Classroom Factors Affecting Students' Self-Evaluations: An Interactional Model
Hermine H. Marshall; Rhona S. Weinstein
Stable URL: http://links.jstor.org/sici?sici=0034-6543%28198423%2954%3A3%3C301%3ACFASSA%3E2.0.CO%3B2-F

Classroom Organization and Student Stratification
Susan J. Rosenholtz; Carl Simpson
Stable URL: http://links.jstor.org/sici?sici=0013-5984%28198409%2985%3A1%3C21%3ACOASS%3E2.0.CO%3B2-N

The Formation of Ability Conceptions: Developmental Trend or Social Construction?
Susan J. Rosenholtz; Carl Simpson
Stable URL: http://links.jstor.org/sici?sici=0034-6543%28198421%2954%3A1%3C31%3AFOACD%3E2.0.CO%3B2-K
The Effect of Classroom Structure on Shared Perceptions of Ability
Susan J. Rosenholtz; Bruce Wilson
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28198021%2917%3A1%3C75%3ATEOCSO%3E2.0.CO%3B2-9

Effects of Ability Grouping on Growth in Academic Achievement
Aage B. Sørensen; Maureen T. Hallinan
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28198624%2923%3A4%3C519%3AEOAGOG%3E2.0.CO%3B2-A