Justice judgments concerning grading in school

Claudia Dalbert, Ulrike Schneidewind, and Annett Saalbach

Martin Luther University of Halle-Wittenberg, Germany


See:
http://www.elsevier.com/wps/find/journaldescription.cws_home/622811/description#description

Corresponding Author:

Prof. Dr. Claudia Dalbert, Department of Educational Psychology, Martin Luther University of Halle-Wittenberg, Franckesche Stiftungen, Haus 5, D-06099 Halle/Saale, Germany; Phone: +49-345-5523811; Fax: +49-345-5527135; e-mail: claudia.dalbert@paedagogik.uni-halle.de
Abstract

This article investigates justice judgments concerning grading. In two studies with secondary students \(N_1 = 350\) academic-track students; \(N_2 = 225\) intermediate \([n = 81]\) and academic-track students \([n = 144]\) attending the 7th to the 12th grade level, participants were presented with vignettes describing three different grading systems: criterion-referenced grading, norm-referenced grading, and individual-referenced grading. The subjects graded and other situational characteristics were varied. In both studies, factor analyses consistently revealed that students' justice judgments were contingent on the comparison standard applied, but independent of other situational characteristics. Criterion-referenced grading was rated as the most just practice, individual-referenced grading as almost just, and norm-referenced grading as almost unjust. The better the grades received the more the students evaluated criterion-referenced grading to be just. Students attending intermediate-track schools did not object to norm-referenced grading as strongly as their peers at academic-track schools did.

Consequences for teachers' grading practices are discussed.

Key words: grading in school, justice, adolescence, school achievement
Justice Judgments Concerning Grading in School

Students and teachers alike are concerned with justice in school. Students want to be treated justly by their teachers, and teachers describe themselves as justice-minded (e.g., Kanders, 2000). Nevertheless, students often complain about being treated unjustly by their teachers. Both the belief that overall events in one's life are just and school-specific experiences of (in)justice have been shown to impact on various domains of the school career. For example, the more justice is experienced in school, the more trust in societal institutions and the less deviant behavior school students report (e.g., Emler & Reicher, 1987; Gouveia-Pereira, Vala, Palmonari, & Rubini, 2004), the better their achievement record (e.g., Dalbert & Stoeber, 2006), and the less distressed they feel in school (e.g., Dalbert & Stoeber, 2005). When asked to describe just and unjust events in school, a majority of students mention their teachers' distributive behavior. Grading, in particular, is a key topic here (e.g., Fan & Chan, 1999; Israelashvill, 1997). In the present study, we therefore investigate students' justice judgments concerning grading in school in more detail.

Grading Systems in School

Types of grading systems. Grading in school implies two processes. First, achievement must be measured. Second, individual achievement must be compared with a given standard (e.g., Rheinberg, 2001). As a rule, three different approaches to grading are distinguished (e.g., Geisinger, Wilson, & Naumann, 1980; Leutner, 1998). In criterion-referenced grading, individual achievement is compared with an objective standard. These objective standards may be defined by the school administrative district, meaning that comparable achievements should be awarded comparable grades across a district, even if a student moves to a new school within the district. In norm-referenced grading, individual achievement is compared with the achievements of a social group – in school, generally with the achievements of one's classmates. This may result in students being awarded lower grades for good performance because the overall achievement in the class was excellent. Moreover, a student moving from
one class to another might get different grades because the standards of the social comparison groups differ. In *individual-referenced grading*, a student's achievement is compared with his or her previous achievement.

In Germany, teachers are supposed to use criterion-referenced grading, and to compare individual achievements with standards defined by the education authorities (Tent, 1998). Often, however, grading practices are in fact based on a mixture of all three approaches (Ingenkamp, 1975). Criterion-referenced and norm-referenced grading, in particular, are not always clearly distinguished (e.g., Flittner, 1985). Norm-referenced grading is most easily applied in the classroom, and may be used to produce a normal distribution of grades within a class. Furthermore, teachers' favorable attitudes towards norm-referenced grading seem to be associated with their belief in the sorting and selection function of grades (Barnes, 1997).

**Influences of grading systems.** Criterion-referenced and norm-referenced grading systems result in individual students being awarded relatively stable grades, provided that they do not move to a new class or to a new school administrative district. Individual-referenced grading, in contrast, is more sensitive to changes in individual achievement and may thus strengthen students' academic motivation (for a review, see Rheinberg, 1987) and learning goal orientations (Dickhäuser & Rheinberg, 2003). Moreover, individual-referenced grading makes it more likely that students will attribute their learning outcomes to variable factors such as effort (Möller & Jerusalem, 1997). Overall, psychological research indicates that individual-referenced grading best supports students' academic success (for a review, see Mischo & Rheinberg, 1995).

**Justice Judgments concerning Grading**

*Equality versus unequal treatment.* Students articulate that they want to be graded equally (e.g., Fan & Chan, 1999; Israelashvill, 1997), Equal treatment would mean the same comparison standard being applied to all students: norm-referenced grading within a specific class, or criterion-referenced grading within a specific school administrative district.
Justice judgments in school

Criterion-referenced grading implies equal treatment for more students (e.g., students from parallel classes) and might thus be seen as more just. Then again, students might consider criterion-referenced grading to be less transparent than norm-referenced grading. It is easy for students to gauge their standing within their own class, but they may not have a clear idea of objective comparison standards, and their knowledge of these objective standards may vary between school subjects and with age. Thus, justice evaluations of criterion-referenced grading may vary between school subjects and across grade levels.

Individual-referenced grading implies unequal treatment because each student is compared with a different standard: his or her own previous achievement. Consequently, it may be the case that students evaluate individual-referenced grading as unjust. Otherwise, while research has shown that unjust grading has detrimental effects on learning outcomes, but also revealed adaptive effects of individual-referenced grading, we might conclude that students evaluate individual-referenced grading as just. In sum, it is far from clear which grading system school students consider to be just. In our study we thus explored students' justice judgments concerning the three different grading systems outlined and investigated whether these judgments differ as a function of the school subject or grade level in question.

*Individual differences in justice judgments.* A justice judgment depicts the individual evaluation of a situation as more or less just. In this view, justice is not a fixed characteristic of a situation; in fact, justice judgments are always subjective (Mikula, 2005). Thorkildsen, Nolen, and Fournier (1994) observed that children with the same educational background exhibit substantial differences in their justice theories concerning educational practices. For example, some students said that effort should be rewarded, whereas others thought that only performance should be rewarded. Consequently, we expected justice judgments concerning grading to vary between students and explored how these differences might be explained.

Personal experiences and dispositions may influence these subjective justice judgments. One important experience with respect to grading is one's own past achievement.
Norm-referenced and individual-referenced grading takes external factors as the teacher behavior and individual variation into account. Attribution research revealed that external attributions as those on teacher behavior and attributions on variable causes as those on varying effort can protect the student self-esteem (Weiner, 1985). Therefore, less successful students may evaluate this kind of "justified" departure from equality to be more just than their better achieving peers. In particular, low achievers might consider individual-referenced grading to be more just, because even small improvements in their performance will be evaluated positively. Thus, we investigated the relationship between justice judgments concerning grading and school achievement.

**Research Questions**

We conducted two studies with the following aims: (a) First, we wanted to find out more about the structure of school students' justice judgments with respect to types of grading systems. We expected students' justice judgments to be contingent on the comparison standard applied, but not on the school subjects considered. In addition, we explored which type of grading system the students consider to be most just. (b) Second, we investigated the relationship between these justice judgments and school achievement. Strong achievers were expected to evaluate criterion-referenced grading as more just than less successful students, who were expected to evaluate individual-referenced grading as more just.

**Schooling in Germany**

German secondary school system consists of three main tracks, of which the "Gymnasium" is the highest. Students attend Gymnasium from grades 5 through 13, completing grade 13 with an "Abitur" diploma that gives them access to university. The "Realschule" is the intermediate-track of the German secondary school system. Students attend Realschule from grades 5 through 10, completing grade 10 with a Realschule diploma. Students who achieve a high grade point average in their Realschule diploma have two options: they may either apply to continue their secondary education and obtain the Abitur at
a Gymnasium, or they may leave school and apply for vocational training in a white-collar job. Students with a low or medium grade point average in their Realschule diploma only have the latter option. Consequently, most Realschule students do not attend university, but go on to work in white-collar jobs. The bottom track of the German secondary school system is the "Hauptschule." Students attend Hauptschule from grades 5 through 9, completing grade 9 with a Hauptschule diploma. Most Hauptschule students go on to work in blue-collar jobs.

In the region of Germany in which our studies took place, the parents are responsible for the assignment of students to secondary school tracks which takes place after the fourth grade; teachers give only advice to the parents. Moreover, within a specific school track, parents can freely select a secondary school; in contrast to primary school, secondary schools are therefore not contingent on the neighborhood.

Study 1

Method

Participants. Three hundred and fifty adolescents attending grades 9 to 12 of three academic-track secondary schools (Gymnasium) in an East German city participated in the first study. Of these, 171 were male and 179 female. Age ranged from 14 to 19 years ($M = 16.9$; $SD = 1.2$). At least two classes were recruited from each school and in each grade level, thus, participants presented about 50% to 70% of the students attending a specific grade level at these schools. The questionnaire was distributed during lesson time. Participants were guaranteed anonymity and volunteered, if their parent gave their permission for participation.

Measures. Regarding achievement, three core subjects are mandatory for all students in German secondary schools: German, mathematics, and one foreign language (usually English). Therefore, as a measure of school achievement, adolescents were asked to indicate the grades they have received in these subjects in their last report. In Germany, school grades range from 1 ("very good") to 6 ("unsatisfactory"), comparable to grades A to F in US schools. To form a general indicator of school achievement, grades were first reversed (so that
a higher value indicates higher achievement) and then averaged across the three subjects (α = .73).

Justice judgments concerning grading were assessed by presenting students four vignettes depicting a grading situation in one sentence. Each vignette was accompanied by three items describing criterion-referenced, norm-referenced, and individual-referenced grading. Thus, justice judgments were assessed by a total of 12 items. Each item was rated on a 6-point scale ranging from 1 ("totally unjust") to 6 ("totally just"). The full instrument as presented to the students is given in Table 1, except the answer scales which were presented in the original directly right-hand behind each item. One vignette depicted a situation in a native language (German) lesson, one a situation in a foreign language (English) lesson, and two described situations in math lessons. Moreover, the vignettes described different grading situations, from an oral test to essay writing.

Results

Students' justice judgments. Principal component analyses were run on the 12 items describing students' justice judgments concerning grading. A three-factor solution was accepted based on the Kaiser-Guttman criteria (λ > 1; Guttman, 1954; eigenvalues: 3.75, 2.90, 1.19, .84) and parallel analysis (Horn, 1965; software: Enzman, 1997). This solution explained 65.3 percent of the variance. A varimax rotation resulted in a simple factor structure defined as a structure with item loadings of > .50 on one factor and < .30 on the other factors. Item statistics are reported in Table 1.

The first factor was characterized by the four items describing individual-referenced grading. These four items were averaged to form a score reflecting the justice judgments concerning individual-referenced grading (α = .86; \( r_{it} \) ranging from \( r_{it} = .66 \) to \( r_{it} = .77 \)). The second factor was characterized by the four items describing norm-referenced grading. These four items were averaged to form a score reflecting the justice judgments concerning norm-referenced grading (α = .80; \( r_{it} \) ranging from \( r_{it} = .56 \) to \( r_{it} = .73 \)). The third factor was
characterized by the four items describing criterion-referenced grading. These four items were averaged to form a score reflecting the *justice judgments concerning criterion-referenced grading* ($\alpha = .78$; $r_{it}$ ranging from $r_{it} = .50$ to $r_{it} = .67$). This pattern of results clearly supports the notion that students based their overall justice judgments concerning grading on the standard of comparison applied. These judgments do not differ across school subjects (German, English, math), between the verbal and math domains, or between different testing situations.

A three-factor ANOVA with the within-subject factor comparison standard (individual-referenced, norm-referenced, criterion-referenced) and gender and grade level (9 to 12) as two between-subject factors revealed significant ($p < .05$) main effects for comparison standard ($F = 211.20; p < .01$) and grade level ($F = 3.20; p = .02$), which were qualified by the interaction of comparison standard and grade level ($F = 2.36; p = .03$). Post-hoc tests (Bonferroni) revealed significant differences between the three justice judgments. Criterion-referenced grading was considered to be the most just approach ($M = 4.47; SD = 0.86$), followed by individual-referenced grading ($M = 3.26; SD = 1.21$), and norm-referenced grading, which was rated as almost unjust ($M = 2.83; SD = 0.96$). The interaction effect was caused by 11th grade students rating individual-referenced grading ($M = 2.93; SD = 1.12$) to be equally unjust as norm-referenced grading ($M = 2.84; SD = 1.01$).

*Achievement and justice judgments.* Correlations between the three justice judgments and grades were inspected. Only the preference for criterion-referenced grading correlated significantly ($p < .05$) with grades ($r = .18; p < .01$), and this relationship persisted when controlled for gender and grade level ($\beta = .20, p < .01$). The better the grades received, the more the students considered criterion-referenced grading to be just.

*Discussion*

One aim of the first study was to investigate the structure and strength of school students' justice judgments concerning grading. The pattern of results indicates that justice
judgments are based on the standard of comparison applied, and not on the school subject. Students evaluated criterion-referenced grading as most just and objected to norm-referenced grading as being almost unjust. In particular, the better students—as defined by their grades—the more they saw criterion-referenced grading as just. Individual-referenced grading was, overall, evaluated as almost just by all, but the 11th graders. We await replication of this unexpected interaction, before speculating why the 11th graders rated the individual-referenced grading as less just than the students from the 9th, 10th, and 12th grade.

We conducted a second study to address some of the limitations of Study 1. Most importantly, Study 1 showed that justice judgments do not differ across school subjects. This may, however, be because two of the three subjects were only mentioned in one vignette. Moreover, judgments may differ if core subjects like math are compared with subjects such as sports. Finally, justice judgments may differ in terms of the grading system applied to young students and students about to leave school (Flittner, 1985). Where young students are concerned, justice judgments may focus on the aspect of nurture and support, and individual-referenced grading might thus be evaluated as most just. When it comes to students about to leave school, however, the selection function of grading may be perceived as more relevant. Consequently, criterion- or norm-referenced grading (e.g., Barnes, 1997) might be seen as more just. Thus, we did a second study to systematically vary two aspects of the grading situation, the school subjects (math versus sport) and the grades of the students to be graded (5th versus final grade).

In the German three-track secondary system, adolescents attending intermediate-track secondary schools tend to show lower intellectual abilities, but also to come from less privileged backgrounds than adolescents attending academic-track schools (Deutsches PISA-Konsortium, 2001). As both student ability and teacher expectations differ markedly across the two tracks, it is important to examine whether the relationships found in Study 1 generalize across the tracks, or whether they are only found for Gymnasium students.
Consequently, students from intermediate- and academic-track secondary schools were recruited for the second study.

Study 2

Method

Participants. Two hundred and twenty-five adolescents attending grades 7, 9, and 12 of two intermediate-track (German: "Sekundarschule"; \( n = 81; \) male = 43, female = 38) and two academic-track secondary schools (German: "Gymnasium"; \( n = 144; \) male = 41, female = 103) participated in the study. One of each academic- and intermediate-track school was recruited in an East German City and one in an adjacent countryside district. The factor city-countryside did not produce any effects, thus, data were collapsed across this factor. Age ranged from 12 to 19 years (\( M = 14.8; SD = 1.9 \)). In the intermediate-track sample, 38 subjects were enrolled in grade 7 and 43 in grade 9. In the academic-track sample, 45 subjects were enrolled in grade 7, 53 in grade 9, and 46 in grade 12. Two classes were recruited from each school and in each grade level, thus, again representing about 50\% to 70\% of the students attending a specific grade level at these schools. The questionnaire was distributed during lesson time. Participants were guaranteed anonymity and volunteered, if their parent gave permission.

Measures. Regarding achievement, grades in German, English, and mathematics were again reversed (so that a higher value indicates higher achievement) and averaged across the three subjects (\( \alpha = .74 \)).

Justice judgments were assessed in a similar format as in Study 1. The questionnaire contained three vignettes describing grades awarded to 5\(^{th} \) graders who had just entered secondary school, and three vignettes describing grades awarded to students in the final class of secondary school (see Table 2). Each vignette depicted the grading situation with one sentence. At each grade level, one vignette referred to math, one to sports (high jump), and one left the subject open. Each vignette was followed by three items describing criterion-
Justice judgments in school

referred, norm-referenced, and individual-referenced grading. Thus, justice judgments were assessed by a total of 6 vignettes and 18 items, and each subject had to answer to all of these 6 vignettes. Vignettes were presented in the same format as in Study 1 with a stem representing the grading situation followed by the three phrases representing the three grading systems. Each item was rated on a 6-point scale ranging from 1 ("totally unjust") to 6 ("totally just"). Finally, in this justice judgment instrument, only the vignettes varied, and not the wording of the grading alternatives. The same phrase depicted a specific grading system after each of the six vignettes. Thus, this justice judgment instrument was of the same format, but more standardized than the one used in Study 1. The full instrument as presented to the students is given in Table 2, except the answer scales which were presented in the original directly right-hand behind each item.

Results

Students’ justice judgments. Principal component analyses were run on the 18 items describing students' justice judgments concerning grading. A three-factor solution was accepted based on the Kaiser-Guttman criteria (λ > 1; Guttman, 1954; eigenvalues: 7.63, 3.40, 1.54, 0.94) and parallel analysis (Horn, 1965; software: Enzman, 1997). This solution explained 70.0 percent of the variance. A varimax rotation resulted in a simple factor structure. Item statistics are reported in Table 2.

The first factor was characterized by the six items describing norm-referenced grading. These six items were averaged to form a score reflecting the justice judgments concerning norm-referenced grading (α = .93; \( r_{it} \) ranging from \( r_{it} = .63 \) to \( r_{it} = .86 \)). The second factor was characterized by the six items describing individual-referenced grading. These six items were averaged to form a score reflecting the justice judgments concerning individual-referenced grading (α = .93; \( r_{it} \) ranging from \( r_{it} = .64 \) to \( r_{it} = .84 \)). The third factor was characterized by the six items describing criterion-referenced grading. These six items were averaged to form a score reflecting the justice judgments concerning criterion-referenced grading.
Justice judgments in school 13

Grading ($\alpha = .87$; $r_{rt}$ ranging from $r_{rt} = .47$ to $r_{rt} = .77$). Again, the pattern of results clearly supports the notion that students based their overall justice judgments concerning grading on the comparison standard applied. This preference did not differ across school subjects (math, sports, no subject specified) or the grade level of the students being evaluated.

A three-factor ANOVA with the within-subject factor comparison standard (individual-referenced, norm-referenced, criterion-referenced), and type of school (intermediate- versus academic-track) and grade level (7, 9, 12) as two between-subject factors revealed significant ($p < .05$) main effects for comparison standard ($F = 99.88; p < .01$), type of school ($F = 4.55; p = .03$), and grade level ($F = 3.91; p = .02$), which were qualified by the interaction comparison standard X grade level ($F = 7.36; p < .01$), the interaction of comparison standard and type of school ($F = 7.32; p < .01$), and the three-way interaction of comparison standard, type of school, and grade level ($F = 6.49; p < .01$).

The interaction of comparison standard and type of school is depicted in Figure 1. Post-hoc tests (Bonferroni) revealed significant differences between the three justice judgments. Both academic- and intermediate-track students considered criterion-referenced grading to be the most just approach (academic-track: $M = 4.84; SD = 1.06$; intermediate track: $M = 4.56; SD = 1.08$). However, intermediate-track students evaluated both norm-referenced and individual-referenced grading as almost just (individual-referenced: $M = 3.96; SD = 1.31$; norm-referenced: $M = 3.72; SD = 1.25$), and thus significantly differed from the academic-track students, who rated norm-referenced grading as more unjust as individual-referenced grading (norm-referenced: $M = 2.83; SD = 1.33$; individual-referenced: $M = 3.42; SD = 1.41$).

Among the intermediate-track students, the judgments of 7th and 9th graders did not differ. Among academic-track students, however, the judgments of 7th graders differed from those of both 9th and 12th graders. More specifically, 7th graders at academic-track schools did not differentiate between criterion-referenced and individual-referenced grading in their
Justice judgments in school

justice ratings (criterion-referenced: $M = 4.42; SD = 1.23$; individual-referenced: $M = 4.18; SD = 1.36$), but rated norm-referenced grading as less just (norm-referenced: $M = 3.49; SD = 1.30$).

Achievement and justice judgments. Correlations between the three justice judgments and grades were inspected. None of the justice judgments correlated with the grades received.

Discussion

The results of the second study again indicate that justice judgments are based on the standard of comparison applied, and not the school subject or the grade level of the student being evaluated. Students rated criterion-referenced grading to be the most just approach. The academic-track students objected to norm-referenced grading as almost unjust, in particular, and rated individual-referenced grading as almost just. The intermediate-track students, in contrast, evaluated both individual-referenced grading and norm-referenced grading as almost just. The youngest academic-track students evaluated individual-referenced grading to be as just as criterion-referenced grading. We again await replication of this unexpected observation before speculating why the 7th grade academic-track students evaluated individual-referenced grading as more just than academic-track students enrolled in grade 9 or 12.

General Discussion

Both studies showed that students have clear opinions about the justice of different grading systems. They base their justice judgments on the comparison standard applied and not on other characteristics of the grading situation. Different school subjects and different grading situations were described in the vignettes, but factor analyses showed that this variation did not explain variance. Departures from criterion-referenced grading system can either be in favor of the student to be evaluated or to that student's disadvantage. Study 1 described individual-referenced grading to the disadvantage of the student to be graded; Study 2 described individual-referenced grading in favor of the student. Norm-referenced grading, by contrast, was formulated to the advantage of the student to be graded in both studies, and
criterion-referenced grading to the disadvantage of the student. In both studies, individual-referenced grading formed an independent factor and did not go along with similar items describing grading in favor or to the disadvantage of the fictional student. This further corroborates the notion that students base their evaluation of teacher grading on the comparison standard applied, independent of other situational characteristics.

A second result is also very clear: Students evaluated criterion-referenced grading as the most just. This held for both studies, and for both academic- and intermediate-track students, and it was relatively independent of school achievement (grades received) and grade level. The application of this objective comparison standard implies that all students in a school administrative district are judged by equal standards. Thus, this pattern of results is in line with the qualitative observations of Fan and Chan (1999) and Israelashvill (1997), who found that students consider equal grading to be just grading. Yet, our studies advance the understanding of equal treatment concerning grading. Equal treatment means the same comparison standard being applied to all students, and this could be norm-referenced grading within a specific class, or criterion-referenced grading within a specific school administrative district. In the eyes of the students, however, just equality is established by criterion-referenced grading only. In fact, norm-referenced grading is considered as less just or even as almost unjust. Moreover, this result may also be interpreted as indicating that teachers are successful in communicating the objective standards that their students have to fulfill in the different subjects.

Students’ evaluations of departures from criterion-referenced grading and hence equality were less consistent than their judgments for this grading system. In particular, academic-track students and intermediate-track students differed with respect to their evaluation of the norm-referenced grading. Whereas academic-track students clearly objected to norm-referenced grading as almost unjust, intermediate-track students evaluated norm-referenced grading as almost just. Thus, we may speculate that the school climate of different
types of school seemed to affect the students' justice judgments concerning norm-referenced grading. More ambitious schools may be more in favor of achievement principles defined by high absolute standards and such a school climate may lead students to favor grading referenced to those high standards. Future studies should more closely examine the overall concept concerning grading at the single school and how this affects the students' justice judgments concerning grading.

The main limitations of the present studies are threefold. (a) We applied a vignette format to assess the justice judgements in which specific grading situations and the three grading possibilities in each situation are characterized in short sentences. These assessments narrow however the participants' perspective on the specific grading domain. It may be that participants have different expectations for types of assignments within domain. Or the justice judgements may be affected by the direction of the deviation from equality, grading to the disadvantage of the student to be graded or grading in favor of the student. Therefore, future studies may systematically vary additional characteristics of the grading situation. (b) School achievement was assessed by school grades received in three core subjects. Clearly, school grades only depict a specific aspect of school achievement. We observed only weak or no associations between the justice judgments and achievement. This may be due to the kind of achievement indicator assessed in both studies. Thus, future study should include other achievement measures and should also consider achievement related dispositions as for example achievement motivation or school related self-efficacy. (c) As these were the first studies on justice judgments concerning different grading systems, the causes of individual differences were not the focus of the present studies. We only included type of school as school characteristic and school grades as an individual characteristic to explain individual differences in justice judgments. Hence, future studies should inquire more closely the school climate and the personal experiences and dispositions which may influence the students' justice judgments. Furthermore, differences in justice judgments concerning grading can be
Justice judgments in school culture-based (Sabbagh, Faher-Aladeen, & Resh, 2004), may depend on the relationship between teacher and students (Thorkildsen, White-McNulty, & Sodonis, 2004), or between the students themselves (Thorkildsen & White-McNulty, 2002). Thus, future studies should also investigate the influence of social factors on justice judgments concerning grading systems. The reliable finding that justice judgments concerning specific grading systems were not contingent on characteristics of the grading situation per se (e.g., the subject) provides a good starting point for this endeavor.

Moreover, justice judgments concerning grading systems are only one aspect of school-based justice evaluations. The interpersonal treatment by the teachers (Schmitt & Montada, 1982), the deviation between received and subjectively deserved grade (Dar & Resh, 2003), or the teachers’ procedural justice (Chory-Assad, 2002) are other dimensions which may as well influence the students’ overall impression to be treated more or less justly at school. To the best of our knowledge, no previous studies have compared the consequences of these different dimensions of the school-based justice evaluations. Future studies should test the unique and complementary impact of these different justice aspects on school career and adolescence development.

If the justice judgments of teachers and their students are in opposition, conflicts may result (Sabbagh, Resh, Mor, & Vanhuysse, 2004), with ensuing negative effects on the students' school careers (Chory-Assad, 2002). Thus, it should be one aim to avoid justice conflicts at school. In this regard, our findings may have potential implications for teachers. Teachers should either strive to apply criterion-referenced grading, because all students evaluated criterion-referenced grading as just, or the teachers should carefully explain the rational behind the application of other grading systems. In addition, teachers may also consider some degree of individual-referenced grading as a possible alternative to criterion-referenced grading. Individual-referenced grading has previously been shown to have positive effects on students' school careers, and our results show that individual-referenced grading
was rated as about just. Thus, in specific situations where good reasons suggest doing so, teachers may opt to conduct individual-referenced grading to boost their students' academic motivation and achievement. However, they should take care not to run into justice conflicts. Future studies should examine in more detail the circumstances under which individual grading is seen as just. Note, as social psychology (e.g., Lerner, 1980) evidenced, not only experienced, but also observed injustice undermines the trust in justice. Thus, the students' evaluation of a just school climate may not only be effected by their experienced grading, but also by observed grading.

Conclusion

Our studies explored school students' justice judgments concerning grading. Students formed their justice judgments based on the comparison standard applied, independent of other characteristics of the grading situation varied in our vignettes. They considered criterion-referenced grading to be most just. Individual-referenced grading was, in general, evaluated as almost just. Norm-referenced grading was, overall, considered as almost unjust. However, students attending less ambitious schools did not object as strongly to norm-referenced grading. As such, teachers in different types of schools should be aware of the justice judgments of their students to avoid justice conflicts and resulting negative effects on the students' school careers. Therefore, future studies should investigate agreement and disagreement between the justice judgments of students and their teachers with respect to grading systems.
References


Justice judgments in school


practices that influence motivation. *Journal of Educational Psychology, 86*, 475-486.


*Psychological Review, 92*, 548-573.
Table 1

*Students' Justice Judgments Concerning Grading: Item Statistics (Study 1)*

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>l1</th>
<th>l2</th>
<th>l3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A student usually does very well in English. In a vocabulary test, she only gets about half of the answers right. What would be a just grade to give her?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because she only got about half of the answers right, she should get a bad grade.</td>
<td>4.54</td>
<td>1.08</td>
<td>.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because her result is still good compared with the rest of the class, she should get a better grade.</td>
<td>2.67</td>
<td>1.21</td>
<td>.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c Because this result is not as good as her earlier work, she should get a worse grade.</td>
<td>3.20</td>
<td>1.43</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A student is called to answer questions in class. Although she is very good at mathematics, she makes a lot of mistakes. What would be a just grade to give her?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because she did well compared with the overall class achievement, she should get a better grade.</td>
<td>2.88</td>
<td>1.12</td>
<td>.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because she did not do as well as in her earlier work, she should get a worse grade.</td>
<td>3.22</td>
<td>1.36</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c Because she made a lot of mistakes, she should get a worse grade.</td>
<td>4.44</td>
<td>1.10</td>
<td>.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A talented student generates a proof in a math test, but his solution is wrong. What would be a just grade to give him?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because the result is not as good as his earlier work, he should get a worse grade.</td>
<td>3.25</td>
<td>1.40</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because not all of the steps in the proof were correct, he should get a worse grade.</td>
<td>4.18</td>
<td>1.18</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c Because his proof is better than that of the others in his class, he should get a better grade.</td>
<td>3.00</td>
<td>1.26</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A good student writes a bad essay in her German class. What would be a just grade to give her?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because she usually writes better essays, she should get a worse grade.</td>
<td>3.38</td>
<td>1.52</td>
<td>.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because her essay is good compared with the overall class achievement, she should get a better grade.</td>
<td>2.71</td>
<td>1.23</td>
<td>.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c Because she has written a bad essay, she should get a worse grade.</td>
<td>4.70</td>
<td>1.07</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Only factor loadings $l > .30$ are given. Item ratings ranged from 1 = "totally unjust" to 6 = "totally just".
Table 2

_Students' Justice Judgments Concerning Grading: Item Statistics (Study 2)_

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>l1</th>
<th>l2</th>
<th>l3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A student in the 5th grade clears a low height in the high jump.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because he/she performed badly, he/she gets a bad grade.</td>
<td>4.45</td>
<td>1.33</td>
<td>.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because his/her performance is still good compared with the class as a whole, he/she gets a better grade.</td>
<td>3.50</td>
<td>1.61</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c Because this result is an improvement on his/her earlier achievements, he/she gets a better grade.</td>
<td>4.13</td>
<td>1.56</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A student in the 5th grade makes some mistakes in a math test.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because he/she performed badly, he/she gets a bad grade.</td>
<td>4.80</td>
<td>1.41</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because his/her performance is still good compared with the class as a whole, he/she gets a better grade.</td>
<td>2.97</td>
<td>1.61</td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c Because this result is an improvement on his/her earlier achievements, he/she gets a better grade.</td>
<td>3.49</td>
<td>1.73</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A student in the 5th grade makes some mistakes in a test.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because he/she performed badly, he/she gets a bad grade.</td>
<td>4.87</td>
<td>1.36</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because his/her performance is still good compared with the class as a whole, he/she gets a better grade.</td>
<td>3.08</td>
<td>1.60</td>
<td>.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c Because this result is an improvement on his/her earlier achievements, he/she gets a better grade.</td>
<td>3.44</td>
<td>1.61</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A student in the final grade clears a low height in the high jump.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because he/she performed badly, he/she gets a bad grade.</td>
<td>4.71</td>
<td>1.38</td>
<td>.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because his/her performance is still good compared with the class as a whole, he/she gets a better grade.</td>
<td>3.15</td>
<td>1.62</td>
<td>.82</td>
<td>.31</td>
<td></td>
</tr>
<tr>
<td>c Because this result is an improvement on his/her earlier achievements, he/she gets a better grade.</td>
<td>3.70</td>
<td>1.60</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A student in the final grade makes some mistakes in a math test.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because he/she performed badly, he/she gets a bad grade.</td>
<td>4.75</td>
<td>1.45</td>
<td>.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because his/her performance is still good compared with the class as a whole, he/she gets a better grade.</td>
<td>3.08</td>
<td>1.57</td>
<td>.80</td>
<td>.35</td>
<td></td>
</tr>
<tr>
<td>c Because this result is an improvement on his/her earlier achievements, he/she gets a better grade.</td>
<td>3.50</td>
<td>1.64</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A student in the final grade makes some mistakes in a test.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Because he/she performed badly, he/she gets a bad grade.</td>
<td>4.80</td>
<td>1.42</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b Because his/her performance is still good compared with the class as a whole, he/she gets a better grade.</td>
<td>3.14</td>
<td>1.57</td>
<td>.82</td>
<td>.32</td>
<td></td>
</tr>
<tr>
<td>c Because this result is an improvement on his/her earlier achievements, he/she gets a better grade.</td>
<td>3.43</td>
<td>1.60</td>
<td>.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Only factor loadings l > .30 are given. Item ratings ranged from 1 = "totally unjust" to 6 = "totally just".
Figure 1. Justice judgements concerning different grading systems for students from academic-track and intermediate-track secondary schools
Justice judgments in school

![Graph showing justice judgments in school]

- **Norm-referenced**
  - Intermediate: 2.83
  - Academic: 3.42

- **Individual-referenced**
  - Intermediate: 3.72
  - Academic: 3.96

- **Criterion-referenced**
  - Intermediate: 4.84
  - Academic: 4.56

The graph illustrates the justice judgments across different reference frameworks, with the y-axis representing perceived justice ranging from very unjust to very just.