Norms and Goals of Appraisal of German Teachers for Students with Learning Disabilities

Jürgen Wilbert*
Matthias Grünke
University of Cologne, Germany

Although there is an ongoing debate about which achievement feedback is most useful, the majority of researchers agree that social comparisons and a focus on competition are inappropriate for students with learning disabilities and otherwise academically-challenged students. They are highly at-risk to be unfavorably influenced by inappropriate responses that teachers give concerning aspects of their performance. Students' achievement and motivation are broadly affected by the type of reference norm that educators use for their feedback. It makes an enormous difference, whether the performance is compared to a learner's former achievements (intra-individual reference norm) or to the achievements of others (social reference norm). The purpose of this study was to investigate the extent to which different kinds of reference norms are used by special education teachers for students with learning disabilities when appraising achievement.

Keywords: Reference Norms; Appraisal; Feedback; Achievement; Motivation; and Learning Disabilities

Most students with learning problems have a personal history of failure in achievement situations. These experiences often result in maladaptive beliefs about the nature of cognitive competencies: intelligence and skills are believed to be innate and static attributes (Dweck, 2008). From this point of view, nothing can be done against one's own failures. Effort is of no avail to change the circumstances; consequently, there is no reason to exert oneself. Perceived absence of control over the outcome of situations is a phenomenon on which Seligman (1975) focused his famous learned helplessness theory on. Together with his colleagues, he discovered that dogs became very pathetic and passive after they were repeatedly confronted with inescapable electric shocks. Eventually, they did not even try to avoid an unpleasant or harmful circumstance to which they had been subjected. In trying to relate their findings to human psychology, Seligman and his workgroup members had to add an important component to their model: the individual's attributional or explanatory style of adverse events. In contrast to dogs, people tend to respond rather differently in the face of negative situations and circumstances. A person may experience the same or a similar adverse event as another, but experience much stronger feelings of depression. If someone internalized a notedly pessimistic explanatory style and views daily hassles as permanent ("This is never going to change."), as personal ("It was my fault."), or as pervasive ("I can't do anything right."), she or he is most likely to develop strong feelings of inadequacy, sadness, and anxiety, as well as an aversion to activity (Peterson, Maier, & Seligman, 1993).

*Please send correspondence to: Jürgen Wilbert, Department of Special Education, University of Cologne, Klosterstr. 79b, 50931 Cologne, Germany. Phone: +49-221-470-2120. Fax: +49-221-470-2148. Email: juergen.wilbert@uni-koeln.de

So, how do these beliefs of helplessness arise in children as they are trying to cope with academic failure? One major reason for these beliefs is the kind of appraisal or feedback that teachers give to students in consequence of their performances (Hattie & Timperley, 2007). It is not primarily the outcome of an action that frustrates students, but the way the environment, and especially the teacher, explains and interprets the cause of it. Depending on this judgment, a negative or low outcome may lead a student to try harder in the future, or it may lead her or him to resign in the light of her or his own insufficiency.

The impact of teachers' appraisal on learning processes has been the focus of educational science for at least 40 years (Brookhart, 2008). Notwithstanding contradictory assumptions of different psychological and educational theories, most researchers share a consensus: Appraisal or feedback has a positive influence on learning and motivation as long as the individual development of a student (temporal comparisons; e.g., "You did better than last time." or "You improved.") is stressed. On the other hand, referring to the achievement of other students (social comparisons; e.g., "You performed in the top quarter of the class.") inhibits learning and motivation processes. This is especially true for extremely low performing children and youth. If the grades of a particular student are worse than those of almost everyone else, downward comparisons are hardly possible. Applying a social reference norm in such a case is eminently counterproductive and inane. But even if downward comparisons were easily possible, they are seldom helpful and motivating in the long run. Gifted students often feel under damaging pressure if the approval of their teachers seems to depend on their performance relative to the learning outcomes of their classmates (Marsh, Trautwein, Ludtke, & Köller, 2008). In any case, using an adequate kind of appraisal or feedback has a greater influence on the academic achievement of children and youth than their prior cognitive abilities or their socioeconomic status (Hattie, 1999). Thus, teachers working with students with learning difficulties should be especially skilled in providing information regarding aspects of their performance in a motivating and encouraging way.

The positive influence of temporal comparisons on learning is a result of the feedback on individual progress that is uncovered for a child or youth. This development is a direct effect of the effort a student invests in accomplishing a task. The nexus between effort and outcome is made visible to the learner. Thus, accomplishment is viewed as dynamic and controllable. The chances are that students take "ownership" of this fundamentally beneficial attitude that connects effort with achievement. On the other side, social comparisons may result in just the opposite process. They put learners into relation to their school fellows. A teacher might, e.g., compare the best and worst students in mathematics in her or his class with each other and thus turn achievement situations into a competitive game with only a few succeeding and many loosing (Elliot & Covington, 2001). Under such circumstances, the self-worth of students is potentially at risk. What makes things worse is the fact that the social position is usually quite robust against changes. A low-achieving individual may improve her or his capacities and skills in the future, but at the same time, her or his class mates improve as well (the high-achieving even more). Even if she or he undertakes more effort for a certain time, this will most likely not increase her or his social position.

All this cuts the linkage between effort and achievement for the learner and increases the probability of developing negative attitudes towards learning.

These types of comparisons emerge from different norms that appraisal refers to. Besides the social and individual (temporal) norm of reference, Heckhausen and his colleagues (Heckhausen Schmalt, & Schneider, 1985) proposed a third one: the criterial norm of reference. Appraisal based on a criterial norm of reference uses task-inherent aspects to rate accomplishments (e.g., "You reached 70 percent of the possible score."). There are comparatively few studies in which the influence of criterial based appraisal on students' motivation and learning has been investigated (Kluger & DeNisi, 1996).

Notwithstanding these insights, it has been repeatedly shown that German teachers from regular schools typically use social comparisons in classrooms (for an overview see Rheinberg & Krug, 1999). However, most of the studies conducted to date have not included teachers from special education classes, and none have focused on teachers for slow learners. Thus, not much is known about their use of different norms of reference when appraising their students. Because of the special needs of children and youth with severe learning problems, the use of individual norms of reference is strongly recommended (Masendorf, 1988). It is still unclear why teachers apply different norms of reference when appraising students' achievements. Studies have previously focused on general personality aspects and general educational goals (Mischo & Groeben, 1995), but those variables explained only small amounts of observed variance. The influence of the context of appraisal situation has not been investigated yet.

PURPOSE AND RESEARCH QUESTIONS

Social psychological research shows that judgment is based on the situational goals that a person strives for (Köpetz, Kruglanski, Xiaoyan, & Orehek, 2008). Teachers' appraisal of achievement may serve different purposes; for example, fostering motivation, establishing or diminishing differences between students, asserting the proficiency, and many others. Some reference norms may be more suitable than others, especially for students who have learning problems. Previous research has indicated that fostering motivation is best attained by an individual norm of reference, changing differences in social ranking by a social norm of reference, and a proficiency-based selection of students by a criterial norm of reference (Köller, 2005). In the light of this reasoning, the reference norm that a teacher of children and youth with learning difficulties uses points to her or his goals of appraisal and can greatly influence student learning. The current study addresses a gap in the research literature regarding reference norms and context of appraisal by asking (1) what kind of reference norms teachers use in working in special education classes for students with learning disabilities and/or other difficulties and (2) how far their use of different norms of reference is influenced by the context of appraisal?

We hold that the context of appraisal has a significant impact on the appliance of the norm of reference. Therefore, we expect a significant interaction of the context of appraisal with each of the norms of reference factors (level of absolute accomplishment, social ranking, and individual development). More specifically, we expect the explained variance of the factor *individual development* to be the highest

in the *fostering motivation* condition, because inciting students to learn is one of the most distinguished goals for special education teachers. Additionally, we expect the explained variance of the factor *social ranking* to be the highest in the *social adaption* condition, and the explained variance of the factor *absolute level of accomplishment* to be highest in the *selection by aptitude* condition. Moreover, we predict the factor *individual development* to be dominant in the *no specific condition*: it is supposed to explain the highest proportion of variance compared to the factors *social ranking* and *level of accomplishment*. Additionally, we prognosticate the pattern of the explained variance of the three factors in the *no specific condition* to be most similar to the pattern of the *fostering motivation* condition, because we expect that special education teachers have a presetting of fostering motivation when given no additional goal of appraisal.

METHOD

Participants and Setting

A sample of 104 teachers of special education classes for students with learning disabilities in Germany (72 females and 32 males) served as subjects. They have been in the teaching field for 10.8 years (*SD* = 9.9) on average. In Germany, children and youth with severe academic problems oftentimes attend special schools for students with learning disabilities. The main criterion for admittance is extensive school failure. Thus, the teachers participating in this study work not only with students with learning disabilities in a narrow sense (i.e., children and youth, who displayed an unexpected difference between their general intellectual ability and their achievement), but also with those students that suffer from severe academic problems including intellectual disabilities and other forms of cognitive difficulties.

Materials and Procedures

The study was conducted using paper and pencil tests. First, participants received instructions on their task, second, a specific context was induced, and third, they performed an appraisal task. At the beginning, all participants were told that they should imagine being a math teacher in three different classes; and their task was to appraise the achievement of their students on the last math test. The context of their feedback was manipulated by giving additional instructions before the presentation of the appraisal task. All participants received the information that their responses would be used as a basis for further decisions concerning the students. Subjects were randomly assigned to three different groups: In the first group, they received the information that students had to be selected for *taking part in an achievement motivation training*. The goal of this program was to foster student achievement based on their individual abilities. In the second group, they were told that students had to be *selected for weekly remedial teaching* in math with the goal of increasing their scores to the class average. Participants in the third group were informed that students were to be *selected for a program of emphasis* in math. The goal was to foster student achievement with respect to their aptitude. The control condition was given no additional information.

We developed an extended version of the *Little Appraisal Task* (Rheinberg 1980; Wilbert & Gerdes, 2009) for measuring the use of different norms of reference. Participants had to appraise the achievement of fictitious students. Therefore, math test scores of 27 learners (three classes each containing nine boys and girls) were displayed (see Table 1). For each student, the scores of the target math test as well as the scores of two proceeding math tests were presented. Participants were instructed that the possible maximum score for each test was 100. The relation of students' achievement within each class was build following the same principle. First, within each class were three types of *individual development* across the three tests: Three students showed increasing scores, three students stagnated, and three students' scores were decreasing. Secondly, within each type of development were three types of *social positions* indicated by the level of the score: high, middle, and low. The difference between the three classes was the *absolute level of performance*. Each student of the middle achieving class scored 17 points above the commensurate student of the low achieving class (i.e., the student with the same individual development and the same social position). The difference between the high and the middle achieving class was 14 points.

Table 1. *The Three Test Scores for All Students of the Three Classes*

Student	Class A			Class B			Class C		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
1	94	90	86	80	76	72	63	59	55
2	80	80	80	66	66	66	49	49	49
3	100	96	92	86	82	78	69	65	61
4	86	86	86	72	72	72	55	55	55
5	84	88	92	70	74	78	53	57	61
6	72	76	80	58	62	66	41	45	49
7	78	82	86	64	68	72	47	51	55
8	92	92	92	78	78	78	61	61	61
9	88	84	80	74	70	66	57	53	49

The appraisal of the accomplishment of each student could be given on a nine-point scale reaching from *very negative* to *very positive*.

Design and Data Analysis

The *context of appraisal* was manipulated between subjects (fostering motivation vs. social adaptation vs. selection by aptitude vs. no specific context). Three factors varied within subjects based on the appraisal of students' achievement: the class *level of accomplishment* (high, middle, low), the *social ranking* of a student within a class (high, middle, low), and the *individual development* of a student (increase,

stagnation, decrease). The multiple independent and dependent variables were analyzed using a multivariate mixed model analysis of variance (MANOVA), performed by SPSS statistical software version 18.0 for Windows. In all statistical tests, $p < 0.05$ was considered significant.

Table 2. *Mean Appraisal for All Students in all Conditions. Scale Points Range From 1 (Negative) to 9 (Positive)*

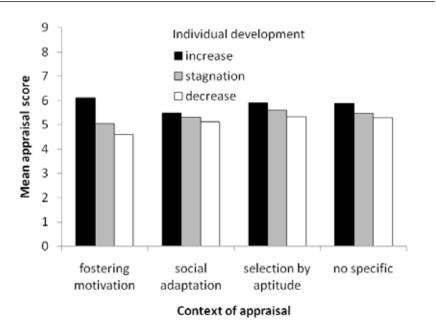| Individual Development and Social Rank	Context of Appraisal			
	Fostering Motivation	Social Adaptation	Selection by Aptitude	No Specific
	High Level of Accomplishment			
Increase				
- High	7.4	7.5	7.2	7.7
- Middle	6.8	6.8	6.3	6.8
- Low	6.4	6.0	5.6	6.4
Stagnation				
- High	6.7	7.4	7.4	7.5
- Middle	5.8	6.5	6.4	6.6
- Low	5.3	6.1	5.8	6.0
Decrease				
- High	5.9	7.5	7.2	7.2
- Middle	5.0	6.5	6.5	6.4
- Low	4.6	5.4	5.5	5.8
	Middle Level of Accomplishment			
Increase				
- High	6.7	6.4	7.0	6.7
- Middle	6.3	5.8	6.3	6.1
- Low	5.9	4.7	5.7	5.4
Stagnation				
- High	5.8	6.3	6.4	6.3
- Middle	5.1	5.5	5.8	5.8
- Low	4.5	4.2	5.0	4.6
Decrease				
- High	4.9	5.9	6.0	6.0
- Middle	5.5	5.6	6.0	5.7
- Low	4.2	4.3	4.5	4.4
	Low Level of Accomplishment			
Increase				
- High	5.8	4.6	5.9	5.3
- Middle	5.3	4.1	5.1	4.7
- Low	4.5	3.4	4.0	4.0
Stagnation				
- High	4.9	5.0	5.5	5.3
- Middle	4.1	3.7	4.5	4.0
- Low	3.5	3.0	3.5	3.2
Decrease				
- High	4.4	4.3	5.1	4.8
- Middle	3.7	3.8	4.2	3.8
- Low	3.3	2.8	3.2	3.5

RESULTS

Descriptive data for the appraisals given by the participating teachers are shown in Table 2. We computed a multivariate mix model analysis of variance (MANOVA). The complete model revealed a significant four-way interaction 4 (context of appraisal) x 3 (level of achievement) x 3 (social ranking) x 3 (individual development) ($F[24, 800] = 1.57$; $MSE = 0.38$; $p < .05$; $\eta_p^2 = .045$). That is, the context of appraisal did influence the pattern of appliance of different norms of reference.

The main effect of the *context of appraisal* was not significant. The *level of achievement* had a significant effect and explained 68% of the variance of appraisal ($F[2, 200] = 212.85$; $MSE = 5.26$; $p < .001$; $\eta_p^2 = .680$). A similarly high amount of variance could be explained by the *social ranking* ($F[2, 200] = 178.77$; $MSE = 2.88$; $p < .001$; $\eta_p^2 = .641$). The impact of the *individual achievement* was also significant, but accounted for considerably less variance ($F[2, 200] = 137.31$; $MSE = 6.98$; $p < .001$; $\eta_p^2 = .164$).

Next, we analyzed the interaction effects of the *context of appraisal* and the three norms of reference factors. The interaction *context of appraisal* x *individual development* was significant ($F[6, 200] = 2.43$; $MSE = 6.98$; $p < .05$; $\eta_p^2 = .068$). Figure 1 reveals that this interaction is due to the stronger impact of the individual development in the *fostering motivation* condition. An a priori contrast of the interaction of the *context of attribution* (fostering motivation vs. the conjoined two other conditions) x *individual development* (stagnation vs. increase) consolidates this impression ($F[1, 100] = 6.64$; $MSE = 12.98$; $p < .05$; $\eta_p^2 = .062$).
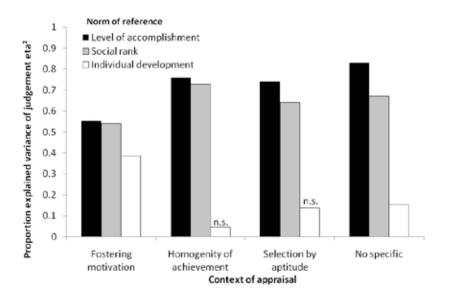
Figure 1. **The impact of the individual development on appraisal separated into the context of appraisal.**

The interaction *context of appraisal* x *social rank* was not significant. Correspondingly, the a priori contrast *context of attribution* (fostering motivation vs. the conjoined two other conditions) x *social rank* (high vs. low) was also statistically irrelevant. The interaction *context of appraisal* x *level of accomplishment* was significant ($F[6, 200] = 3.14$; $MSE = 5.26$; $p < .01$; $\eta_p^2 = .086$), but the a priori contrast of the interaction of the *context of attribution* (selection by aptitude vs. the conjoined two other conditions) x *level of accomplishment* (high vs. low) was not ($F < 1$). Hence, the context of appraisal had an impact on the influence of the *level of accomplishment*, but this influence was not as we hypothesized. Post hoc contrasts revealed that the level of accomplishment was significantly less influential in the *fostering motivation* condition compared to the conjoined to other conditions ($F[1, 100] = 4.73$; $MSE = 8.23$; $p < .05$; $\eta_p^2 = .045$).

Next, we analyzed the overall use of norms of reference when no specific context of appraisal is induced. Figure 2 displays the explained variance of the three norms of reference for each context of appraisal. As indicated in Figure 2, students' individual development clearly had the lowest impact on appraisal when no specific context was induced. The pattern in the no specific context condition is most different when compared to the *fostering motivation* condition and most similar to the *social adaptation* and *selection by aptitude* condition.

Figure 2. The explained variance of appraisal by the three norms of reference split into the different contexts of appraisal. All but the two marked bars (n.s.) show a significant influence on appraisal.

Discussion

The purpose of this study was to elicit the norms of reference used by special educators when giving feedback to students with learning disabilities and to analyze the relationship between the context of appraisal and the application of specific reference norms. There were several findings from the results that further understanding of how special education teachers provide feedback to students. First, teachers generally did not stress individual development and did not provide feedback that is geared towards enhancing motivation, if they are given no specific assignment before having to rate the performance of students in a test. Second, it was shown that the aforementioned factors (context of the assessment situation and the appliance of specific reference norms) are significantly associated with each other. In accordance with the postulated hypotheses, teachers focus on individual development when appraising students' achievement with the purpose of fostering their motivation. Third, neither the class level of accomplishment, nor the social ranking of individual learners was considered by special educators in the proposed manner when rating achievements in order to select candidates for remedial courses based on their relative proficiency level or their aptitude.

There are surprisingly few recent studies that have systematically investigated the kind of appraisal that special educators use in their classrooms. This paper can certainly be considered as an important contribution to shedding light on this essential topic. However, our results are very general in nature. The kind of appraisal and feedback that teachers give usually depends on a variety of factors. In an early study, Dweck, Davidson, Nelson, and Enna (1978) were able to demonstrate that educators normally use an individual reference norm more often with boys than with girls when providing learners with information about a poor performance. Burnet and Stuart (2001) showed that the effect that different kinds of appraisals have on students varies with age: During their time in elementary school, children's need for effort feedback declines while the need for ability feedback increases. Depending on students' ages, teachers usually respond to this concern by providing them with different kinds of information regarding a given aspect of their performance. Ethnicity seems to also play an important role in this context (De Luque & Sommer, 2000). It can be assumed that the participants would have rated the performances of the fictitious students in some cases differently, had they been provided with more information about them (gender, age, ethnic background, etc.). If the feedback had not been related to the achievement of learners in a math test, but to some other curriculum-related outcome, this study could have also yielded slightly different results. The same applies to a scenario where teachers had to appraise the performance of actual and familiar students, instead of fictitious and unknown students, because personal relationships seem to also play an important role in this context (Santangelo, 2007). However, this study would have hardly been feasible, had we confronted our subjects with a tangled mass of information about 27 students.

As indicated above, teachers working with students with learning difficulties are under a special obligation to be well-informed about the beneficial and damaging power of appraisal and feedback in order to teach them effectively. And they must implement this knowledge in their classrooms. Because if they portrayed only insufficient expertise in this regard, they could cause much more harm than if they were

exclusively teaching typically achieving students. Against the background of these circumstances, the results of this study give cause for concern. Even though many aspects about the influence of motivating appraisal are still underresearched and even though the validity of the findings in this paper is limited by several factors, one assertion appears to be reliable: Especially when teachers are working with students with learning difficulties, it is crucial to point out to them any improvements in their achievements in order to keep them motivated to learn. To repeatedly compare their relatively poor performance with the more favorable learning outcomes of other students usually decreases their willingness to try harder next time. Thus, there is a great need to provide teachers—working with students with learning disabilities and/or difficulties—with substantiated and ample information about the effects different kinds of feedback have on them and their performance in the classroom.

References

Brookhart, S. M. (2008). *How to give effective feedback to your students*. Alexandria, VA: ASCD.

Burnett, P. C., & Stuart, C. U. (2001). Elementary students' preferences for teacher praise. *Journal of Classroom Interaction, 36*, 16-23.

De Luque, M. F., & Sommer, S. M. (2000). The impact of culture on feed-backing-seeking behavior. *Academy of Management Review, 25*, 829-849.

Dweck, C. S. (2008). Self-theories and lessons for giftedness: A reflective conversation. In T. Balchin, B. Hymer, & D. J. Matthews (Eds.), *The Routledge international companion to gifted education* (pp. 308-316). New York: Routledge.

Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness. *Developmental Psychology, 14*, 268-276.

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review, 13*, 73-92.

Hattie, J. (1999). *Influences on student learning*. Retrieved January 10, 2010, from http://www.education.auckland.ac.nz/webdav/site/education/shared/hattie/docs/influences-on-student-learning.pdf

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81-112.

Heckhausen, H., Schmalt, H. D., & Schneider, K. (1985). *Achievement motivation in perspective*. New York: Academic Press.

Kluger, A. N. & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254-284.

Köller, O. (2005). Formative assessment in classrooms: A review of the empirical German literature. In Centre of Organisational Research and Instruction (Ed.), *Formative assessment: Improving learning in secondary classrooms* (pp. 265-279). Paris: Organisation for Economic Co-operation and Development (OECD).

Köpetz, C. E., Kruglanski, A. W., Xiaoyan, C., & Orehek, E. (2008). Goal systemic effects in the context of choice and social judgement. *Social and Personality Psychology Compass, 2*, 2071-2089.

Marsh, H. W., Trautwein, U., Ludtke, O., Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs. *Journal of Educational Psychology, 100*, 510-524.

Masendorf, F. (1988). Alternativen für eine sonderpädagogische Beurteilungs- und Notenpraxis. [Alternatives for special education teachers to assign marks and appraise]. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete, 57*, 269-276.

Mischo, C., & Groeben, N. (1995). Bezugsnormorientierung: Warum sich LehrerInnen unterscheiden. [Reference norm orientation: Why teachers differ]. *Empirische Pädagogik, 9*, 423-459.

Peterson, C., Maier, S., and Seligman, M. E. P. (1993). *Learned helplessness: A theory for the age of personal control*. New York: Oxford University Press.

Rheinberg, F., & Krug, S. (1999). *Motivationsförderung im Schulalltag*. [Fostering motivation in school]. Göttingen: Hogrefe.

Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation*. [Achievement appraisal and lerning motivation]. Göttingen: Hogrefe.

Santangelo, T. (2007, November). *What makes a great teacher? A unique, multi-dimensional application of content analysis to capture student voice*. Paper presented at the 16th Annual World Congress on Learning Disabilities, Marlborough, Massachusetts (USA).

Seligman, M.E.P. (1975). *Helplessness: On depression, development, and death*. San Francisco, CA: W. H. Freeman.

Wilbert, J., & Gerdes, H. (2009). Die Bezugsnormwahl bei der Bewertung schulischer Leistungen durch angehende Lehrkräfte des Förderschwerpunktes Lernen. [The choice of reference norm when appraising students' achievement of prospective teachers for students with learning disabilities]. *Heilpädagogische Forschung, 35,* 122-135.