Big fish in big ponds: A multilevel analysis of test anxiety and achievement in special gifted classes

THOMAS GOETZ[1], FRANZIS PRECKEL[2], MOSHE ZEIDNER[3], & ESTHER SCHLEYER[3]

[1]University of Konstanz, Germany; Thurgau University of Teacher Education, Switzerland; [2]University of Trier, Germany & [3]Center for Interdisciplinary Research on Emotions, University of Haifa, Haifa

Abstract
This study analyzes the effects of individual achievement and achievement level of student reference group on test anxiety in a national sample of 769 gifted Israeli students (grade levels 4–9), which was previously investigated by Zeidner and Schleyer (1999a). We hypothesized that when controlling for individual achievement, students' experiences of test anxiety should increase with the increasing ability level of their peer reference group. It was assumed that this effect was largely mediated by reference group effects on academic self-concept (big-fish-little-pond effect). Zeidner and Schleyer found that gifted students within a gifted peer reference group showed higher levels of test anxiety than gifted students within a non-gifted peer reference group. Of note, the present study focused exclusively on gifted students attending special gifted classes. The main research question was whether or not the assumed effects of individual and class achievement can be found for gifted students in special gifted classes when taking the variance of achievement level (grades) of the special gifted classes into account. Using hierarchical linear modeling (HLM) methodology, the assumed effects were vindicated for this special group of high ability students. Thus, in line with previous results, the Worry component of test anxiety was more highly reactive to the effects of individual achievement than the Emotionality component. Also, in line with our theoretical assumptions, achievement/anxiety relations were largely mediated by the effects of academic self-concept.

Keywords: *Test anxiety, worry, emotionality, academic self-concept, achievement, big-fish-little-pond effect, gifted education*

Numerous studies have investigated the complex pattern of relations between achievement and test anxiety (Zeidner, 2007). Test anxiety has been found to interfere with achievement both in laboratory settings as well as in true-to-real life testing situations in school or collegiate settings (see Zeidner, 1998, for review). While the vast majority of these studies analyzed relations between students' experiences of anxiety and their *individual achievement*, little research has been conducted on achievement-related peer reference group effects on test anxiety. The present study investigated achievement/test anxiety relations by taking both individual achievement as well as the *achievement level of the peer reference group* into account. In the next section, we give a brief general description of the nature of the performance/anxiety relationship, and discuss the role of academic self-concept as a

Correspondence: Thomas Goetz, Faculty of Humanities, Universitaetsstr. 10, D-78457 Konstanz, Germany. E-mail: thomas.goetz@uni-konstanz.de

possible mediator of this relationship. Further, we concentrate on gifted samples and outline findings on peer reference group effects on test anxiety by referring to two recent studies. In one study by Zeidner and Schleyer (1999a), gifted students in full time homogenous classes are compared with gifted students who attend regular classes and participate in a one-day pull-out program. The other study by Preckel, Zeidner, Goetz and Schleyer (in press) focuses on students in full time homogenous classes exclusively.

A good number of studies, including integrative reviews and meta-analyses, have demonstrated clear academic benefits of special homogenous classes for gifted students (Goldring, 1990; Kulik & Kulik, 1982; Rogers, 1993). This does not come as a surprise when one takes into account that in most heterogeneous classroom settings over 80% of the time is spent on whole class activities, with no attention to differentiating for the gifted (Archambault, Westberg, Brown, Hallmark, Emmons, & Zhang, 1993). However, Goldring (1990) reports that when grouped homogenously, gifted students may develop feelings of incompetence and anxiety as a reaction to the increased ability and performance of the reference group.

The nature of the performance-anxiety relationship is best viewed as reciprocal in nature. High levels of test anxiety, with elevated levels of worry and cognitive interferences, absorb part of the capacity needed for attention, short-term memory, problem-solving, or other cognitive processes required for successful completion of the cognitive task at hand (Baddeley, 1986). Test anxiety also produces certain aversive patterns of motivation, coping, and task strategies that interfere with learning and performance. Poor performance outcomes lead to further anxiety over time, generating a vicious circle of increasing anxiety and degrading performance (Wells & Matthews, 1994).

Theoretical models focusing on the impact of performance on test anxiety often refer to cognitive appraisals as mediating this relationship. Among the different constructs addressed by appraisal theories of emotions (Scherer, Schorr, & Johnstone, 2001), subjective control is one of the most relevant (Spielberger & Vagg, 1995). Following this line of research, the control-value theory of Pekrun (2006) on achievement emotions explicitly assumes that achievement/test anxiety relations are mediated by academic self-concept. To our knowledge, this mediation assumption has yet to be empirically tested. However, the empirical conditions for such a mediation are given (Kenny, Kashy, & Bolger, 1998), namely, significant relations between the predictor and the dependent variable (achievement/test anxiety relations; see above), significant relations between the predictor and the mediator (achievement/self-concept relations; Valentine, DuBois, & Cooper, 2004), as well as significant relations between the mediator and the dependent variable (self-concept/test anxiety relations; Hembree, 1988). The present study sets out to test the hypothesis stating that the relationship between achievement and test anxiety is mediated by academic self-concept.

Numerous studies have attested to the bi-dimensional nature of the test anxiety construct. A widely accepted definition of test anxiety proposed by Spielberger construes it as an individual's disposition to react with extensive worry, intrusive thoughts, mental disorganization, tension, and physiological arousal when exposed to evaluative situations (Spielberger & Vagg, 1995). Concerning the differentiation of the Worry and Emotionality components of test anxiety, advanced by Liebert and Morris (1967), anxiety is viewed as including a cognitive (Worry) and affective (Emotionality) component. Specifically, Worry is viewed primarily as cognitive concern about the consequences of failure, whereas Emotionality is defined as consisting of perceptions of autonomic reactions evoked by evaluative stress. These two components are revealed to be empirically distinct, though

correlated (Liebert & Morris, 1967). A bulk of studies has shown that the Worry component is more strongly related to academic achievement than the Emotionality component (see meta-analyses of Hembree, 1988 and Seipp, 1991) as it absorbs much more cognitive resources than the Emotionality component. Due to the differential relationships of Worry and Emotionality with performance, we separate these constructs for the analyses in the present study. In addition to assessing the nexus of relations between individual achievement and both Worry and Emotionality, we also sought to determine the relations between these two anxiety components with the achievement level of the reference group. This is a substantially new contribution to the field of differentiating Worry and Emotionality in test anxiety research. It appears that from both a theoretical and applied perspective, it is important to determine to what degree the achievement level of the reference group impacts both cognitive and affective components of test anxiety. Educators, for example, may want to know whether or not students' test anxiety, evoked by the high achievement level of the reference group, is primarily characterized by concerns about the consequences of failure (Worry) or by autonomic reactions due to evaluative stress (Emotionality). Knowledge on the anticipated effects of the reference group on Worry and Emotionality might help educators to deal more adequately with students facets of test anxiety (e.g., by explicitly avoiding focusing on consequences of test performance or by trying to establish a class climate which reduces evaluative stress reactions).

As mentioned above, one shortcoming of the studies analyzing achievement/test anxiety relations is that they have primarily concentrated on *individual achievement* outcomes. In other words, they have analyzed whether students showing different levels of individual achievement also differ in their experiences of test anxiety. There are very few studies which took the *achievement level of the peer reference group* as a predictor of individual levels of test anxiety into account. A prominent theoretical approach referring to the effects of the achievement level of the peer reference group on individual experiences of self-concept is the "big-fish-little-pond effect" (BFLPE; Marsh, 1987) which has been repeatedly attested to in gifted student populations when comparing gifted students attending special gifted classes with those attending regular classes (e.g., Craven, Marsh, & Print, 2000; Marsh, Chessor, Craven, & Roche, 1995; Zeidner & Schleyer, 1999a).

As academic self-concept can be assumed being a mediator of the achievement/test anxiety relations, the BFLPE on academic self-concept might help to explain the achievement/test anxiety relations. The BFLPE hypothesizes that it is better for academic self-concept, and, consequently, for test anxiety to be a big fish in a little pond (i.e., good student in a peer reference group of average ability), than to be a small fish in a big pond (i.e., good student in a peer reference group of high ability). Marsh and Parker (1984) proposed a social frame of reference model to explain the BFLPE. According to this model, self-perceptions in educational settings are largely shaped by the process of social comparison. With increasing ability level of a peer reference group (e.g., class), students often compare themselves with high ability peers and are compared by their teachers with more intellectually able peers, which, in turn, affects the feedback (e.g., grades) students receive. Thus, the likelihood for upward social comparisons with more able students, as well as the likelihood to get lower grades, increases with the ability level of the reference group. This, in turn, results in lower academic self-perceptions (for overviews, see Koeller, 2004; Marsh, 2005; Marsh & Craven, 2002), and might consequently result in higher levels of test anxiety.

Zeidner and Schleyer (1999a) reported data demonstrating reference group effects on test anxiety. In their study, focusing mainly on gifted students, Zeidner and Schleyer

(1999a) used different educational contexts, i.e., regular versus special gifted classes, to operationally define the ability level of the respective reference group. In a recent study, Preckel, Zeidner, Goetz, and Schleyer (in press) showed that reference group effects also appear in special gifted classes, exclusively, when taking the achievement level of the peer reference group (average achievement outcomes of students within classes) into account.

In their study, Zeidner and Schleyer (1999a)[1] compared the level of test anxiety and academic self-concept reported by gifted students in regular classes with the respective levels reported by their gifted counterparts in special gifted classes. The working assumptions and testable hypotheses of this study were based on the BFLPE (Marsh, 1987; see above). As expected, gifted students in special gifted classes (i.e., big fish in big ponds) reported higher levels of test anxiety and lower levels of academic self-concept than gifted students in regular classes (i.e., big fish in little ponds). Zeidner and Schleyer also investigated the impact of the peer reference group on the two components of test anxiety (Worry, Emotionality; Liebert & Morris, 1967) separately. Their analyses showed that gifted students in special classes (high achievement level of peer reference group) and mixed classes (medium achievement level of peer reference group) were significantly more differentiated on the Emotionality than on the Worry component. In other words, their results indicated that the Emotionality component was more reactive to the achievement level of the peer reference group than the Worry component.

Preckel et al. (in press) investigated whether the BFLPE on academic self-concept also holds *within* the group of special gifted classes. They operationally defined the achievement level of the reference group by average achievement outcomes of students within classes. The level of achievement was sufficiently variable in various gifted classes to allow for peer reference group effects for students of very high ability. By using hierarchical linear modeling (HLM) methodology, Preckel et al. found a clear BFLPE on academic self-concept within the group of special gifted classes: When controlling for the positive achievement/self-concept relations on the individual level, average class achievement was negatively related to individual levels of academic self-concept.

Zeidner and Schleyer (1999a) found negative consequences of high achieving reference groups for test anxiety and academic self-concept when looking at the individual level of analysis, and contrasting the levels of both constructs in gifted students attending regular classes (little ponds) and gifted students attending special gifted classes (big ponds). Preckel et al. (in press) showed in subsequent analyses that the negative effects of the achievement level of the reference group on academic self-concept also holds *within* the group of special gifted classes. However, Preckel et al. did not investigate the effects of individual achievement and achievement level of the peer reference group on *test anxiety* within the group of special gifted classes. An obvious assumption following the results of both studies is that the negative consequences of high achieving contexts for *test anxiety* also hold *within* the group of special gifted classes. Testing this assumption was the primary aim of the present study. A number of subsidiary goals of this study are addressed in the following paragraph.

Zeidner and Schleyer's results indicated that the Emotionality component was more reactive to the peer reference group than the Worry component. In view of this finding, a second aim of the present study was to investigate whether this finding can be replicated within the group of special gifted classes. Further, if there are positive consequences of individual achievement and negative consequences of the achievement level of the reference group for test anxiety or its two components, these effects can be assumed to be *mediated* by academic self-concept. That is to say, when controlling for the presumed strong impact of academic self-concept on students' test anxiety levels, there might be no additional

association between achievement and test anxiety—both on an individual as well as group (class) level (Kenny et al., 1998). This assumption can be derived from appraisal theories of anxiety, and particularly from the control-value theory of Pekrun (2006). Testing this mediation assumption within the group of special gifted classes was the third aim of the present study.

Research Questions and Hypotheses

Based on empirical results relating to the effects of the achievement level of the reference group on test anxiety, as reported in the literature by Zeidner and Schleyer (1999a) and Preckel et al. (in press), assumptions of appraisal theories of anxiety and Pekrun's (2006) control-value theory on achievement emotions, the following hypotheses were tested:

(1) When simultaneously considering both individual achievement outcome and average achievement level of the peer reference group (class), individual achievement outcome will be negatively related to individual test anxiety. Also, the average achievement level of the class will be found to be positively related to the individual level of test anxiety.

(2) The effects of individual achievement and average class achievement will be found to be different for the Worry and Emotionality components of test anxiety. Specifically, in line with cumulative empirical evidence, the Worry component of test anxiety will be shown to be more reactive to individual achievement outcomes than the Emotionality component (see meta-analyses of Hembree, 1988 and Seipp, 1991). Concerning the achievement level of the reference group, the Emotionality component will be shown to be more reactive to the achievement level of the peer reference group (class) than the Worry component. We note that this hypothesis is exploratory in that there is no theoretical work bearing on the differential effects of the achievement level of the peer reference group on different test anxiety components. Further, there are no empirical findings beyond the results reported by Zeidner and Schleyer (1999a).

(3) Achievement is predicted to work through the mediating effects of self-concept in impacting on test anxiety. Thus, when taking the effects of academic self-concept on test anxiety into account, no additional effects of achievement (individual and class level) on test anxiety will be observed.

Method

Participants and Data Collection

The sample of this study consists of 789 intellectually gifted Israeli students in 37 special homogeneous classes for gifted students. Placement test scores, originally used for screening students for special programs for the gifted, were made available from the Ministry of Education for a subset of the sample ($n = 482$). The gifted students had median percentile scores of $M = 90.19$ ($SD = 11.86$). Due to the HLM approach we adopted in this study, classes including less than 10 students were not integrated in our multi-level analyses (Felson & Reed, 1986). This selection criteria resulted in analyzing 33 classes with a medium class size of about 23 students ($M = 22.62$; $SD = 5.56$). Of the 769 students (31% female) within these 33 classes, 330 students were drawn from elementary schools (grades 4–6), and 439 students were drawn from junior high schools (grades 7–9).

Gender and individual differences data were collected via self-report instruments that were group-administered to students during regular classroom period. Students' grades were obtained directly from students' school records.

Variables and Measures

Test anxiety was assessed by a 12-item version of the Hebrew adaptation (TAI/H; Zeidner, Nevo, & Lipschitz, 1988) of Spielberger's (1980) *Test Anxiety* Inventory (TAI). The two components of Worry (e.g., "I worry about possible failure when studying for an exam") and Emotionality (e.g., "When I'm taking an exam, I feel uncomfortable and tense") were gauged by six items each. Students rated items on a four-point Likert-type scale (1 = almost never, 4 = almost always). The TAI/H total scale, comprised of a linear combination of the 12 items comprising the Worry and Emotionality subscales, showed satisfactory internal consistency ($\alpha = 0.86$; $M = 28.23$, $SD = 7.45$), as did the Worry ($\alpha = 0.75$; $M = 12.64$, $SD = 3.91$) and Emotionality ($\alpha = 0.83$; $M = 15.58$, $SD = 4.31$) components.

Academic self-concept was assessed by a 12-item Likert-type scale patterned after selected items from the experimental Hebrew translation (Zeidner, 1995) of the Academic Self-Concept subscale of Bracken's Multidimensional Self Concept Scale[2] (MSCS; Bracken, 1992). Students evaluated themselves on a four-point scale (1 = strongly disagree, 4 = strongly agree) on items gauging various facets of learning, academic competence and scholastic performance (e.g., "I am bright and learn things easily"). The reliability of this scale was $\alpha = 0.85$ ($M = 28.23$, $SD = 7.45$).

Scholastic achievement was indexed by the following six measures: students' end of year grades in three cardinal school subjects (Mathematics, English Language, Biblical Literature) gathered for two consecutive years. Grades for individual students were gleaned from student records. Grades were based on the conventional numeric grading scale employed in Israeli schools, ranging from 0 to 100 in each school subject, with higher values indicating better achievement. Passing scores, per subject, are normally set at 50, with 70 serving as the average score. Grades in cardinal subjects serve as the major basis for evaluating student achievement, and serve as pivotal criteria for student placement procedures and within-school streaming in math and language. Scores across all required subjects during each school year are used to determine student's grade point average. Reliability of scholastic achievement was $\alpha = 0.81$ ($M = 526.09$, $SD = 41.12$).

Unfortunately, there are no current empirical data available bearing on the external validity of teacher-assigned grades in the Israeli school system against external criteria, such as standardized achievement test scores or international assessment results (Personal communication, Avital Dorman, Israeli Academy of Sciences, 10 June 2007). Also, there are no current data available related to the between-school variance of teacher grades for the subjects reported in this study. It is noted that when teacher grades in the upper high school levels are compared to external matriculation score grades, the former are more restricted and show less variance (ranging from 60 to 95 compared to about 40–100 for matriculation grades).[3]

In this regard, it is important to note that studies on the BFLPE typically use standardized achievement tests to measure both individual student achievement and class-average achievement. However, the BFLPE can also be found when using students' grades for operationally defining individual student achievement and class-average achievement (Preckel et al., in press). As grades are a more salient source of feedback for students than standardized test scores (Marsh, 1987; Wylie, 1979), course grades can even be assumed to have stronger effects on academic self-concepts. Evidence for this theoretical assumption was found for the individual level in a recent study by Marsh, Trautwein, Luedtke, Koeller, & Baumert (2005). However, concerning the effects of the average class achievement on academic self-concept, we lack empirical knowledge of how the impact of class achievement operationally defined by scores of standardized achievement tests differs from the impact of class achievement operationally defined by school grades.

Data Analyses

Preliminary analysis. We first report zero-order correlations among study measures. This analysis is important to see whether intercorrelations of independent variables were sufficiently low to exclude multicollinearity.

Multilevel analyses. As we wanted to model the effects of individual data and data aggregated across groups (average achievement levels in classes) simultaneously, we used multilevel analysis (HLM; Raudenbush & Bryk, 2002). The analyses were implemented via HLM™ 6.01 (Raudenbush, Bryk, Cheong, & Congdon, 2001).

For each dependent variable, namely test anxiety (TAI/H total scale), as well as its two components (Worry and Emotionality), three models were computed. Independent variables were cumulatively aggregated in our models. Integrating independent variables, step by step, allowed us to investigate how the inclusion of the variable influences the effects of previously included independent variables on the outcome variable.

Model 1 included individual achievement (Level 1 predictor) and aggregated class achievement (Level 2 predictor) as independent variables. This model should show whether there are different effects of individual achievement and achievement level of the reference group on our dependent variables when taking both achievement indicators simultaneously into account. Model 2 added gender (Level 1 predictor) to the predictor stock in our analysis. We integrated gender as a control variable in our models because previous research has shown that females report significantly higher levels of test anxiety than males (Hembree, 1988; Seipp & Schwarzer, 1996; Zeidner, 1998). Further, gender group differences have been shown to be of considerably greater magnitude on the Emotionality than on the Worry component of test anxiety (Zeidner & Nevo, 1992). Model 3 added academic self-concept (Level 1 predictor) to our HLM analysis. The important question for all of our three dependent variables (i.e., total test anxiety total score, Worry, Emotionality) is whether the effects of achievement, both on individual and class level, change when controlling for the effects of academic self-concept on the dependent variable. Significantly reducing the effects of achievement on the dependent variable would support the mediation assumption (see Kenny et al., 1998), namely that achievement-anxiety relations are mediated by academic self-concept.

Results

Preliminary analyses

Table I presents zero-order product-moment correlations among all of our measures. Intercorrelations of independent measures were sufficiently low (rs ≤ 0.53) to exclude multicollinearity.

Multilevel analyses[4]

Test anxiety (total score) as dependent variable. Model 1 (see Table II) shows that when taking achievement on an individual and class level simultaneously into account, individual achievement was significantly *negatively* ($\beta = -0.16$) related to test anxiety, while mean class achievement was significantly *positively* ($\beta = 0.13$) related to test anxiety. This result supports the assumed differential effects of individual achievement vs. reference group achievement level on test anxiety. Both effects are still significant even when controlling for

Table I. Zero-order correlation coefficients among measures.

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Test anxiety total score	–					
(2) Worry component	0.90**	–				
(3) Emotionality component	0.91**	0.63**	–			
(4) Academic self-concept	−0.37**	−0.38**	−0.28**	–		
(5) Achievement	−0.19**	−0.26**	−0.09**	0.53**	–	
(6) Gender	−0.08**	−0.03	−0.12**	0.02	−0.06*	–

Notes. The test anxiety total score comprises of a linear combination of its Worry and Emotionality component. Gender was scaled as follows: females $=0$, males $=1$. $n=769$ students (31% female). $*p<0.05$. $**p<0.01$.

gender effects (Model 2). When integrating academic self-concept as a Level 1 predictor (individual level) in the analysis (Model 3), we found a significantly negative effect of this variable on test anxiety[5] ($\beta = -0.37$). The significant effects of achievement found in Models 1 and 2 vanished both at the individual, as well as at the class level. This result indicates that the achievement-test anxiety relation was mediated by academic self-concept.[6]

Table II. Results of the three multi-level-analyses: effects of general achievement score, gender and academic self-concept on either test anxiety, Worry, and Emotionality.

	Model 1		Model 2		Model 3	
	β	p	β	p	β	p
Test anxiety						
Individual level:						
Achievement	−0.16	<0.01	−0.20	<0.01	0.00	0.98
Gender			−0.10	0.01	−0.09	0.03
Academic self-concept					−0.37	<0.01
Class level:						
Achievement	0.13	0.03	0.15	0.02	0.06	0.33
Worry						
Individual level:						
Achievement	−0.21	<0.01	−0.23	<0.01	−0.05	0.29
Gender			−0.03	0.41	−0.02	0.69
Academic self-concept					−0.32	<0.01
Class level:						
Achievement	0.10	0.11	0.11	0.09	0.03	0.68
Emotionality						
Individual level:						
Achievement	−0.09	0.05	−0.13	<0.01	0.05	0.35
Gender			−0.14	<0.01	−0.13	<0.01
Academic self-concept					−0.34	<0.01
Class level:						
Achievement	0.13	0.02	0.15	0.01	0.08	0.16

Notes. All outcome and predictor variables on individual and class level were standardized ($M=0$, $SD=1$) across the entire sample, so that unstandardized beta weights which are reported in the table correspond to standardized beta weights. Gender was scaled as follows (before standardization): females $=0$, males $=1$. Concerning the significance of the βs, exact p-values are depicted in the case that they are not below 0.01.

Worry component as dependent variable. Model 1 (see Table II) shows that individual achievement was significant and negatively ($\beta = -0.21$) related to Worry, while the relation between mean class achievement and Worry did not reach statistical significance. Controlling for gender effects did not change this pattern of results (Model 2). When integrating academic self-concept as a Level 1 predictor (individual level) in the analysis (Model 3), we found a significantly negative effect of this variable on Worry[5] ($\beta = -0.32$). The significant effect of achievement on the individual level that was found in Models 1 and 2 vanished. This result indicates that achievement/Worry relation was mediated by academic self-concept.[6]

Emotionality component as dependent variable. Model 1 (see Table II) shows that individual achievement is *negatively* ($\beta = -0.09$; $p = 0.05$) related to Emotionality, whereas mean class achievement is significantly *positively* related to the Emotionality component of test anxiety ($\beta = 0.13$). This result supports the assumed different effects of individual achievement and achievement level of the reference group on Emotionality. This relational pattern remains when controlling for gender effects (Model 2). When integrating academic self-concept as a Level 1 predictor (individual level) in the analysis (Model 3), we found a significantly negative effect of this variable on Emotionality[5] ($\beta = -0.34$). The significant effects of achievement on an individual, as well as class level, found in Models 1 and 2 vanished. This result suggests that the achievement-Emotionality relation was mediated by academic self-concept.[6]

Discussion

In the present study, we investigated the effects of individual achievement and achievement level of the reference group on test anxiety for gifted students attending special gifted classes. The achievement level of the peer reference group was operationally defined by average school grades within class. In line with our assumptions, we found the following effects on test anxiety: individual achievement was negatively related to test anxiety while the average achievement of the class was positively related to test anxiety. While Zeidner and Schleyer (1999a) found negative consequences of the achievement level of the peer reference group on test anxiety when defining the achievement level of the reference group categorically (regular classes vs. special gifted classes), the results of our study show that the consequences of high achieving reference groups on test anxiety could also be found *within* the group of the special gifted classes when differentiating them according to their achievement level. The replication of the negative consequences of the achievement level of the reference group on test anxiety *within* special gifted classes indicates that for high ability students, the achievement level of the peer reference group plays a crucial role in the formation of test anxiety. In other words, there seems to be nothing like a threshold which "immunizes" against peer reference group effects on test anxiety. Upward social comparisons also seem to play an important role for gifted students. From an educational perspective, educators, school governance officers and politicians should be aware of these results indicating that students in gifted classes are prone to peer reference group effects (see also Preckel et al., in press).

When analyzing the two components of test anxiety separately, our results on the individual level are in line with numerous other studies (Hembree, 1988; Seipp, 1991). Accordingly, the Worry component of test anxiety was more reactive to individual achievement outcomes than the Emotionality component. Concerning the effects of the achievement level of class on Worry and Emotionality, our results do not corroborate Zeidner and Schleyer's findings. While they found the Emotionality component to be more

reactive to peer reference group, in our study we found the effects of the achievement level of class to be nearly the same in terms of effect size for Worry and Emotionality. Of note, the effects were significant for Emotionality, exclusively, indicating that the Emotionality component is more reactive to peer reference group. However, before drawing conclusions this finding needs to be replicated in further studies.

In line with assumptions derived from appraisal theories of anxiety, and particularly Pekrun's (2006) control-value theory of achievement emotions, we found that the achievement/test anxiety relations were mediated by academic self-concept. When taking the effects of academic self-concept on test anxiety into account, we found no additional effect of individual or class level achievement on test anxiety (c.f., Kenny et al., 1998). In other words, the "BFLPE on test anxiety" found in our study was in fact a "classical" BFLPE on academic self-concept with subsequent effects of self-concept on experiences of test anxiety. Self-concept as a mediator of the achievement/anxiety relations was also found both for the Worry, as well as for the Emotionality component, of test anxiety. It is important to note that our study was cross-sectional in nature, and that the results found in our study cannot be interpreted in terms of causality (cf. Cole & Maxwell, 2003). Thus, the achievement/self-concept/test anxiety relations and their underlying mechanisms, as found in our study, should be replicated by using a longitudinal design.

We concentrated in our study on the BFLPE exclusively. However, there is a further important effect on academic self-concept based on group affiliation with assumingly subsequent consequences on test anxiety which was not at the center of our study, namely the "basking-in-reflected-glory" effect (Cialdini, Borden, Thorne, Walker, Freeman, & Sloan, 1976). This effect is based on the assumption that self-concept may be enhanced by membership in groups which are positively valued through the reflected glory of accomplishments or good qualities of other group members. In the present study, students attending special gifted classes might "bask in reflected glory", which in turn, might lead to an enhancement of self-concept, and, consequently, to a reduction of test anxiety (e.g., "If I am smart enough to be in this special gifted class with all these other very intelligent students, then I must be very smart"). It has often been shown that both of these counter-balancing effects, i.e., the BFLPE and the "basking-in-reflected-glory" effect, are likely to affect self-concept simultaneously (e.g., Marsh, Kong, & Hau, 2000; Marsh, Koeller, & Baumert, 2001; McFarland & Buehler, 1995). Thus, the typically observed negative BFLPE, as found in our study, can be assumed to be actually a net effect, with the BFLPE evidenced to be stronger than the "basking-in-reflected-glory" effect. In other words, the "basking-in-reflected-glory" effect might have operated even though it was overshadowed by the BFLPE. Consequently, it can be assumed that when controlling in our study for the "basking-in-reflected-glory" effect, including relevant variables in the model might result in a stronger, that is, more negative BFLPE. Further studies on the effects of achievement on test anxiety mediated by academic self-concept may additionally take the "basking-in-reflected-glory" effect into account.

A further limitation of our study is that we used teacher assigned school grades to operationally define students' achievement, both at individual and class level. While numerous studies have used school grades as achievement indicators on the individual level, with one exception, namely the Preckel et al. (in press) study, we do not know studies using school grades as indicators of the class average achievement. However, class average achievement is a salient indicator for individuals in view of judging their own achievement in relation to those of their classmates. Consequently, it is warranted to use grades as indicators of class average achievement when carrying out analysis on the BFLPE. Due to

the grading-on-a-curve phenomenon, variance in the average class achievement based on grades might be smaller than the variance in average class achievement based on standardized achievement tests (c.f., Hodge & Coladarci, 1989). This can first of all be assumed when classes across different school types (due to streaming, tracking or ability grouping), and, consequently, strongly varying ability levels are investigated. Even levels of achievement as investigated by standardized achievement tests might strongly vary across classes of different school types, grades can be assumed to be relatively similar (Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). From a statistical perspective, this reduced variance might reduce the impact of class achievement on self-concept when average class grades are used as achievement indicators. However, in the present study, students did not attend different school types, and there was a relatively strong and significant ($p<0.01$) variance in the average class achievement based on grades. Both aspects lead to the assumption that using grades for operationally defining average class achievement has been appropriate in our study. Future studies might use both the results of standardized achievement tests *and* school grades to operationally define class average achievement, in order to investigate the potential different strengths of both indicators.

Our findings on the effects of mean class achievement on self-concept and test anxiety for gifted students in special classes are in line with the findings of studies comparing gifted students in special classes with gifted students in regular classes. Delcourt et al. (1994) found that students in special classes attained higher levels of achievement than their peers in regular classes, but they conversely had "the lowest perception of their academic competence, preference for challenging tasks, sense of acceptance by peers, internal orientation, and attitudes to learning" (p. 77). In addition, teacher expectations are usually higher for students in gifted classes. This combination of high teacher expectations and students' low perceptions of their abilities makes it even more important that instruction in gifted classes supports the students in developing a realistic and positive self-concept. Our results indicate that enhancing students' individual academic self-concept might reduce their test anxiety. This is of high practical concern since test anxiety not only negatively affects achievement, but also subjective well-being (Diener, 2000). Furthermore, test anxiety can be assumed to relate negatively to student involvement in terms of course enrollment, career aspirations, and occupational choices in specific domains (cf., Wigfield, Battle, Keller, & Eccles, 2002).

There are numerous ways of reducing test anxiety of students in a direct way. An intervention can target test anxiety itself (emotion-oriented treatment; e.g., using relaxation techniques); the control appraisals underlying emotions (appraisal-oriented treatment; e.g., cognitive restructuring); the academic competencies determining students' agency (competence-oriented treatment; e.g., training of learning skills); and the educational environment, including classroom instruction (design of academic environments). A more complete analysis of test anxiety treatment is outlined in Zeidner (1998, chapters 13–16). Further, Goetz, Frenzel, Pekrun, and Hall (2006) give examples for regulating test anxiety derived from theories on "emotional intelligence." In addition, our results suggest an indirect way of reducing students' test anxiety by enhancing their academic self-concept. Studies have investigated various possibilities to counter the negative effects of social comparisons which are assumed to be the main reason for the BFLPE on academic self-perceptions. Probably the most promising approach is to encourage teachers to use individual frames of reference when providing students with achievement-related feedback (instead of a social or achievement-criteria/peer group related frame of reference). An individual frame of reference is likely to reduce the

amount of students' social comparisons of individual achievement outcomes with those of other students in class. Such an individual frame of reference employed by teachers has indeed been found to have positive effects on students' academic self-perceptions (Koeller, 2004). Consequently, it might be assumed that an individual frame of reference can reduce the negative consequences of an increasing level of class achievement on students' experiences of test anxiety. We found a single study that looked at the effects of individual frame of reference (ipsative teacher feedback) on the BFLPE (Luedtke & Koeller, 2002). In line with results reported by Koeller (2004), Luedtke and Koeller found that employing an individual frame of reference in class showed a positive effect on students' academic self-concept. However, this study could not confirm that the BFLPE was smaller in classes where teachers employed a strong individual frame of reference. Clearly, further research on this topic is warranted.

Notes

1 In addition to the results outlined by Zeidner and Schleyer (1999a), these authors report results derived from the same data set in three further publications (Zeidner & Schleyer, 1999b,c,d). These analyses are based on further research and educational perspectives than those in the Zeidner and Schleyer (1999a) study. However, even some of the findings reported there are partly related to the present study, we exclusively refer to Zeidner and Schleyer (1999a) due to the fact that there are no results in the Zeidner and Schleyer (1999b,c,d) studies which would contribute to the present research questions beyond the Zeidner and Schleyer (1999a) study.

2 Bracken (1992) reported a high internal consistency of the academic self-concept subscale ($\alpha = 0.91$; range: 0.88–0.92 across grade levels 5–12; $n = 2510$; 52% females). With respect to the validity of the academic self-concept scale, Bracken found a strong correlation with general school self-concept ($r = 0.66$) and weak correlations with the self-concept of opposite-sex relations ($r = 0.18$) or same-sex relations ($r = 0.31$)—each assessed by the respective scale of Marsh's (1990) Self-Description Questionnaire–II.

3 We thank Professor Sorel Cahn of the Hebrew University and Avital Dorman of the Israeli Academy of Sciences for providing us with very helpful information on teacher-assigned grades in Israeli school.

4 For the purposes of HLM analyses, all outcome and predictor variables on both the individual and class level were standardized ($M = 0$, $SD = 1$) across the entire sample (see Aiken & West, 1991). Fully unconditional models which are equivalent to a one-way analysis of variance with classes as a random factor indicated that there was significant ($p < 0.01$) between-class variance in our variables. The intraclass correlation coefficients (ICCs) were 0.19 for academic achievement, 0.16 for test-anxiety (total score), 0.14 for Emotionality, and 0.18 for Worry. Unstandardized beta weights are reported in the results section. As all outcome and predictor variables were standardized, these unstandardized beta weights correspond to standardized beta weights for all variables. All of our Level 1 and Level 2 predictors were uncentered (what is due to our standardization procedure equivalent to being centered around the grand mean). As there was no reason to assume varying slopes across our classes, we calculated constant slope models which make estimations more efficient and the models more parsimony (Snijders & Bosker, 1999).Our HLM specifications for the three models were as follows (ACH = individual scholastic achievement; ACH_AGG = aggregated class achievement; ACAD = academic self-concept):

Model 1: $y_{ij} = \gamma_{00} + \gamma_{01}*\text{ACH\_AGG}_{\cdot j} + \gamma_{10}*\text{ACH}_{ij} + u_{0j} + r_{ij}$

Model 2: $y_{ij} = \gamma_{00} + \gamma_{01}*\text{ACH\_AGG}_{\cdot j} + \gamma_{10}*\text{ACH}_{ij} + \gamma_{20}*\text{GENDER}_{ij} + u_{0j} + r_{ij}$

Model 3: $y_{ij} = \gamma_{00} + \gamma_{01}*\text{ACH\_AGG}_{\cdot j} + \gamma_{10}*\text{ACH}_{ij} + \gamma_{20}*\text{GENDER}_{ij} + \gamma_{30}*\text{ACAD}_{ij} + u_{0j} + r_{ij}$

5 Preckel et al. (in press) analyzed the effects of achievement (individual and class level) on academic self-concept referring to the same sample as studied in the present study. The effect was $\beta = 0.57$ ($p < 0.01$) on the individual level and $\beta = -0.22$ ($p < 0.01$) on class level, reflecting the "classical" BFLPE on academic self-concept (Marsh, 1987).

6 The three models were run a second time while controlling for the effects of class membership in either elementary school or junior high school on test anxiety (total score), and its Worry and its Emotionality components, respectively. This was achieved by including a dummy-coded variable which coded this membership at the individual level. As inclusion of this variable did not change the pattern of results, only the results for the analyses without this dummy variable are reported.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Archambault, F. X. Jr., Westberg, K. L., Brown, S. W., Hallmark, B. W., Emmons, C. L., & Zhang, W. (1993). *Regular classroom practices with gifted students: Results of a national survey of classroom teachers*. Storrs, CT: National Research Center on the Gifted and Talented, University of Connecticut.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Bracken, B. A. (1992). *Multidimensional self-concept scale*. Austin, TX: Proed.

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, *34*, 366–375.

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*, 558–577.

Craven, R. G., Marsh, H. W., & Print, M. (2000). Gifted, streamed, and mixed-ability programs for gifted students: Impact on self-concept, motivation, and achievement. *Australian Journal of Education*, *44*, 51–75.

Delcourt, M. A. B., Loyd, B. H., Cornell, D. G., & Goldberg, M. D. (1994). *Evaluation of the effects of programming arrangements on student learning outcomes*. Virginia: The National Research Center on the Gifted and Talented.

Diener, E. (2000). Subjective well-being. *American Psychologist*, *55*, 34–43.

Felson, R. B., & Reed, M. D. (1986). Reference groups and self-appraisals of academic ability and performance. *Social Psychology Quarterly*, *49*, 103–109.

Goetz, T., Frenzel, A., Pekrun, R., & Hall, N. C. (2006). Emotional intelligence in the context of learning and achievement. In R. Schulze, & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 233–253). Cambridge, MA: Hogrefe & Huber Publishers.

Goldring, E. B. (1990). Assessing the status of information on classroom organisational frameworks for gifted students. *Journal of Educational Research*, *83*, 313–326.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, *58*, 7–77.

Hodge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement. A review of the literature. *Review of Educational Research*, *59*, 297–313.

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* Vol. 1 (pp. 233–265). Boston, MA: McGraw-Hill.

Koeller, O. (2004). *Konsequenzen von Leistungsgruppierungen* [Consequences of homogenous groups with regard to school performance]. Münster, Germany: Waxmann.

Kulik, J. A., & Kulik, C.-L. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal*, *19*, 415–428.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, *20*, 975–978.

Luedtke, O., & Koeller, O. (2002). Individuelle Bezugsnormorientierung und soziale Vergleiche im Mathematikunterricht. Einfluss unterschiedlicher Referenzrahmen auf das fachspezifische Selbstkonzept der Begabung [Individual reference norm and social comparisons in mathematics classes: The impact of different frames of reference on the domain-specific self-concept of ability]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *34*, 156–166.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*, 280–295.

Marsh, H. W. (1990). *Self-Description Questionnaire-II*. San Antonio, TX: Psychological Corporation.

Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *German Journal of Educational Psychology*, *19*, 119–127.

Marsh, H. W., & Parker, J. W. (1984). Determinants of self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well. *Journal of Personality and Social Psychology*, *47*, 213–231.

Marsh, H. W., & Craven, R. (2002). The pivotal role of frames of reference in academic self-concept formation: The big-fish-little-pond effect. In F. Pajares, & T. Urdan (Eds.), *Adolescence and education* Vol. 2 (pp. 83–123). Greenwich, CT: Information Age.

Marsh, H. W., Chessor, D., Craven, R., & Roche, L. (1995). The effect of gifted and talented programs on academic self-concept: The big fish strikes again. *American Educational Research Journal*, *32*, 285–319.

Marsh, H. W., Kong, C. K., & Hau, K. T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong. *Journal of Personality and Social Psychology*, *78*, 337–349.

Marsh, H., Koeller, O., & Baumert, J. (2001). Reunification of East and West German school systems: Longitudinal multilevel modeling study of the big-fish-little-pond effect on academic self-concept. *American Educational Research Journal*, *38*, 321–350.

Marsh, H. W., Trautwein, U., Luedtke, O., Koeller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*, 397–416.

McFarland, C, & Buehler, R. (1995). Collective self-esteem as a moderator of the frog-pond effect in reactions to performance feedback. *Journal of Personality and Social Psychology*, *68*, 1005–1070.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*, 315–341.

Preckel, F., Zeidner, M., & Goetz, T, Schleyer, E. (in press). Female 'big fish' swimming against the tide: The 'big-fish-little-pond effect' and gender ratio in special gifted classes. *Contemporary Educational Psychology*.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (second edition). Thousand Oaks, CA: Sage Publications.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2001). *HLMTM 5. Hierarchical linear and nonlinear modeling*. Lincolnwood: Scientific Software International.

Rogers, K. B. (1993). Grouping the gifted and talented: Questions and answers. *Roeper Review*, *16*, 8–12.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds). (2001). *Appraisal processes in emotion*. Oxford, UK: Oxford University Press.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research*, *4*, 27–41.

Seipp, B., & Schwarzer, C. (1996). Cross-cultural anxiety research: A review. In C. Schwarzer, & M. Zeidner (Eds.), *Stress, anxiety, and coping in academic settings* (pp. 13–68). Tuebingen, Germany: Francke-Verlag.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.

Spielberger, C. D. (1980). *Test anxiety inventory: Preliminary professional manual*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., & Vagg, P. R. (1995). Test anxiety: A transactional process. In C. D. Spielberger, & P. R. Vagg (Eds.), *Test anxiety: Theory, assessment, and treatment* (pp. 3–14). Washington, DC: Taylor & Francis.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, *98*, 788–806.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relations between self-beliefs and academic achievement: A systematic review. *Educational Psychologist*, *39*, 111–133.

Wells, A., & Matthews, G. (1994). *Attention and emotion: A clinical perspective*. Hove, UK: Lawrence Erlbaum.

Wigfield, A., Battle, A., Keller, L. B., & Eccles, J. S. (2002). Sex differences in motivation, self-concept, career aspiration, and career choice: Implications for cognitive development. In R. De Lisi, & A. McGillicuddy-De Lisi (Eds.), *The development of sex differences in cognition* (pp. 93–124). Westport, CT: Ablex Publishing.

Wylie, R. C. (1979). *The self-concept* Vol. 2. Lincoln, NA: University of Nebraska Press.

Zeidner, M. (1995). *Hebrew version of Bracken's Multidimensional Self-Concept Scale* (experimental edition). University of Haifa: Author.

Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Plenum Press.

Zeidner, M. (2007). Test anxiety: Conceptions, findings, conclusions. In P. Schutz, & R. Pekrun (Eds.), *Emotion in education* (pp. 165–184). San Diego, CA: Academic Press.

Zeidner, M., & Nevo, B. (1992). Test anxiety in examinees in a college admission testing situation: Incidence, dimensionality, and cognitive correlates. In K. A. Hagtvet, & B. T. Johnsen (Eds.), *Advances in test anxiety research* (Vol. 7 pp. 288–303). Lisse, The Netherlands: Swets & Zeitlinger.

Zeidner, M., & Schleyer, E. J. (1999a). The big-fish-little-pond effect for academic self concept, test anxiety, and school grades in gifted children. *Contemporary Educational Psychology*, *24*, 305–329.

Zeidner, M., & Schleyer, E. J. (1999b). Test anxiety in intellectually gifted school students. *Anxiety, Stress and Coping*, *12*, 163–189.

Zeidner, M., & Schleyer, E. J. (1999c). The effects of educational context on individual differences variables, self-perceptions of giftedness, and school attitudes in gifted adolescents. *Journal of Youth and Adolescence*, *28*, 687–703.

Zeidner, M., & Schleyer, E. J. (1999d). Evaluating the effects of full-time vs. part-time educational programs for the gifted: Affective outcomes and policy considerations. *Evaluation and Program Planning*, *22*, 413–427.

Zeidner, M., Nevo, B., & Lipschitz, H. (1988). *The Hebrew version of the test anxiety inventory*. Haifa: University of Haifa.