The Significance of Self-Reported Anxious Symptoms in First-Grade Children

Nick Ialongo,[1,4] Gail Edelsohn,[2] Lisa Werthamer-Larsson,[1] Lisa Crockett,[3] and Sheppard Kellam[1]

It is generally believed that prior to the middle to late elementary school years children's reports of anxious symptoms represent nothing more than transient developmental phenomena. In light of the limited empirical study of this issue and its import to the allocation of mental health resources, the present study seeks to provide empirical evidence of the significance of anxious symptoms in children younger than 7. Specifically, utilizing an epidemiologically defined population of 1197 first-grade children, followed longitudinally from the fall to spring of first grade, we examine the stability, prevalence and caseness of children's self-reports of anxious symptoms. Self-reported anxious symptoms proved relatively stable over 4-month test–retest interval. In addition, they appeared to have a significant impact on academic functioning in terms of reading achievement. These findings on stability, caseness, and prevalence suggest children's self-reported anxious symptoms in the early elementary school years may have clinical significance. However, further study is necessary before firm conclusions can be drawn.

Benjamin, Costello, and Warren (1990) noted the limited epidemiologic study of anxiety as a symptom, syndrome or disorder in children and suggested one reason for this may be the belief that childhood anxiety symptoms are transient and innocuous (see Lapouse & Monk, 1959). The lack of epidemiologic study is particularly true for children less than 7 or 8.

Indeed, only a handful of studies of anxious symptoms and disorders in community samples have included children younger than 8 and even fewer included children less than 7. Moreover, in only two of these studies was the study population truly epidemiologically defined and not a sample of convenience (Bird et al., 1988; Fleming, Offord, & Boyle, 1989). Furthermore, child self-reports of anxiety were not the basis for selection into the sample in either study. Rather, parent reports were the gateway to child interviews and while highly specific, parent reports tend to fall short in terms of sensitivity (Angold, 1988).

Serving to highlight the importance of epidemiologic studies is the fact that the selection biases associated with the use of clinic samples and other samples of convenience are avoided (Berkson, 1946). Thus, generalization to known populations can be drawn and incidence and prevalence rates can be validly inferred (Kellam, 1990). Given the relative absence of epidemiologic studies of anxiety in young children, informed decisions with regard to the allocation of the limited treatment and preventive intervention services available are difficult to make (Kellam, 1990, National Institute of Mental Health, 1991). In light of the limited empirical study of this; issue and its import to the allocation of mental health resources, the present study seeks to provide empirical evidence of the significance of anxious symptoms in children younger than 7. Specifically, in line with Quay (1986), and the criteria he set for judging the validity of a putative psychiatric syndrome, we examine the stability of first graders' self-reports of anxious symptoms over an interval of 4 months. We also examine their "caseness" (Beardslee, Klerman, Keller, Lavori, & Podorefsky, 1985), that is, the degree to which anxious symptoms impact upon adaptive functioning in the domains of academic achievement, peer relations, attention/concentration, social participation, and acceptance of authority. With respect to the issue of caseness, there are a number of ways in which anxiety may have detrimental effects on cognitive functioning and social development. For example, the concentration problems that often characterize anxiety disorders and syndromes may serve to disrupt the mastering of new, or complex, academic problems. In terms of the detrimental effects on social development, normal interactions with peers may be precluded as a result of the social withdrawal often seen in anxious children.

Besides assessing the significance of anxious symptoms in terms of their stability and association with adaptive functioning, we also sought to obtain an estimate of the prevalence of severe anxiety in the population of first graders studied. An understanding of the extent of a mental health problem is necessary for the appropriate allocation of resources to treat or prevent the problem. Our estimate of the prevalence of severe anxiety was based on the administration of the Revised Children's Manifest Anxiety

Scale (Reynolds & Richmond, 1985) to the population studied and the use of the clinical cutoffs reported by Reynolds and Richmond. A critical issue also addressed in this paper is the extent to which the reliability, caseness, and prevalence of anxious symptoms vary by gender. Our examination of gender differences is consistent with the findings from a number of studies of internalizing symptoms in young children. For example, in their large scale study of depressive symptoms in 8- to 13-year-old boys and girls, Lefkowitz and Tesiny (1985) reported a larger number and a greater strength of relations between depressive symptoms and indices of social functioning in girls as opposed to boys. Finally, in accord with the framework outlined by Kellam (1990), wherein psychological well being is seen as both a consequence of, as well as an antecedent to, adaptation to developmental tasks, we examined the impact of adaptation to developmentally relevant tasks on psychological well-being in terms of the manifestation of anxious symptoms. For school-aged children, these tasks include academic achievement, social participation with teachers and classmates, acceptance of authority, and sustained attention and concentration in the classroom.

METHOD

Subjects

Participants were 1,197 first-grade children from 19 public elementary schools in the eastern catchment area of the city of Baltimore, Maryland. The schools were selected from each of five sociodemographically distinct areas in eastern Baltimore. The children were originally recruited for participation in two school-based, preventive intervention trials targeting early learning and aggression. Special education and gifted classrooms were excluded from the pool of potential classrooms in light of the fact the preventive interventions targeted regular or mainstream classrooms. Children had been randomly assigned to classrooms prior to assignment of classrooms to intervention conditions. Schools were randomly assigned to either an intervention or control condition within a geographic area.

The five geographic areas in which the participating schools are located were defined by census tract data and vital statistics from the Baltimore City Planning Office. These areas vary by ethnicity, type of housing, family structure, income, unemployment, violent crime, suicide, and school dropout rates. With regard to the gender, ethnicity, and age of the subject population, 49.1% was male, 65.6% African–American, 31.6% white, 0.3 Asian, 1.0% Native American, 0.3% Hispanic, and for 1.2% of the children, ethnicity was either missing or refused. At first grade, the average age of

the children was 6.6 years (SD ± 0.48). Of the 1,197 children available for participation in fall of first grade, written parental consent was obtained for 1,002 children, or 83.8% of the population. Sixty-one or 5.1% of the children's parents or guardians refused to allow their children to participate, whereas 28 gave verbal consent. Eight children had transferred out of the participating schools prior to consent being requested, whereas 94 parents or guardians failed to respond to the consent request. Chi-square analyses revealed that refusal rates varied as a function of geographic area $[\chi^2(16, N = 1,197) = 43.67, p = .0002]$. The highest rates of refusal were in Areas 1 and 4, which are made up primarily of middle-income, two-parent families, living in well-maintained row or detached homes.

Given our primary interest in this paper in describing the natural or untreated course of anxious symptoms in school-aged children, only the 684 control subjects were included in the analyses described below. Importantly, there were no differences between the control and treated subjects in terms of gender, age, ethnicity, geographic area, or consent/participation rates. The exclusive use of the control subjects allows us to generalize our results to similarly defined, untreated community populations.

Assessment Design and Measures

The design of the preventive trial from which these data were drawn included child assessments in the fall and spring of first grade. The fall of first-grade assessments were carried out prior to the initiation of the interventions. The data gathered in these assessments included children's self-reports of anxious symptoms, standardized achievement scores, peer nominations of likability and social participation, and teacher ratings of children's concentration problems, social participation, and aggressive behavior. With the exception of the standardized achievement scores, the assessment data were generally collected from teachers and children within the same $2\frac{1}{2}$-hour period.

Revised Children's Manifest Anxiety Scale (R-CMAS; Reynolds and Richmond, 1985). The R-CMAS is a 37-item, self-report instrument designed to assess the level and nature of anxiety in children and adolescents from 6 to 19 years old. The child responds to each statement by marking a *yes* or *no* answer. A response of *yes* indicates that the item is descriptive of the child's feelings or actions, whereas a response of *no* indicates that the item is generally not descriptive. The *yes* responses are counted to determine a total anxiety score. Based on the data presented by Reynolds and Richmond from their standardization sample of children from 6 to 19 years of age, the R-CMAS has demonstrated good internal consistency and

there is evidence supporting construct validity (Reynolds and Richmond, 1985). For the total anxiety score, the median reliability estimate was .82 (range .79 to .85). Stability coefficients for the total anxiety score ranged from .98 for a 3-week interval to .68 for a 9-month interval (Reynolds and Richmond, 1985).

As there is considerable variation in the reading skills of elementary school-aged children, the format for children to record their answers on the R-CMAS was slightly modified. Simple symbols, that is, pictures of common shapes (circle and square) and objects (e.g., ball, apple, etc.), were used to indicate answer choices and to show the place of each item on the answer sheet. More specifically, to help them find the correct place on the answer sheet to indicate their answers, the children were asked to put their fingers on the picture of the object corresponding to the appropriate item on the answer sheet. Then children were asked to indicate *yes* or *no* by placing an X over the circle or square next to their fingers, respectively. For example, for the first item children were asked to place their fingers on the "ball" on their answer sheet. The item was then read aloud twice by the interviewer to the child. The child was then asked to indicate whether their answers were *yes* or *no*. If their answers were *yes*, they were instructed to draw an X in the circle. If they were *no* they were instructed to draw an X in the square. Previous administration procedures have relied on either the child reading the items and circling *yes* or *no* in response to a particular item or having an examiner read the items while the child follows along and circles his/her choices. No other changes were made in the administration format or the content of the R-CMAS.

Teacher Observation of Classroom Adaptation—Revised (TOCA-R; Werthamer-Larsson, Kellam, & Wheeler, 1991). The TOCA-R is a structured interview with the teacher, administered by a trained assessor who follows a script precisely and responds in a standardized way to issues the teacher initiates. The interviewer records the teacher's ratings of individual children as the interview proceeds. Teachers respond to 36 items pertaining to the child's adaptation over the last 3 weeks to classroom task demands. Adaptation is rated by teachers on a 6-point scale and covers three factor analytically derived dimensions: concentration problems, authority–acceptance/aggression, and shy behavior/social participation. Werthamer-Larsson *et al.,* reported test–retest correlation over a 4-month interval with different interviewers of .60 or higher for each of these factors and coefficient alphas of .96, .92 and .85, respectively. Scores on the concentration problems factor correlated .44 with overall performance on the California Achievement Test. In addition, a correlation of .67 was found between ratings on the authority–acceptance/aggression factor and peer nominations of aggression.

Finally, the shy behavior/social participation factor correlated moderately with peer nominations of likability ($r = -.34$).

The Peer Assessment Instrument (PAI). The PAI is a modified version of the Revised Pupil Evaluation Inventory (R-PET; Pekarik, Prinz, Leibert, Weintraub, & Neale, 1976). Ten items were selected from the R-PEI for use in the PAI. In some cases, items were shortened and reworded in order to be more readily comprehended by beginning first graders. The constructs of interest in the present study were social participation/shy behavior and likability. The items used to assess social participation include "plays alone a lot" and "too shy to make friends," whereas the likability construct was made up of the following items: "are liked by everyone," "are your best friends," and "have very few friends." In terms of administration, a descriptor is read aloud to the class and the children are then instructed to circle the picture(s) of those children who fit the descriptor. Raw scores on each of the above dimensions are converted to standard scores based on the distribution of nominations within a child's classroom.

The California Achievement Test (CAT; Forms E & F). The CAT represents one of the most frequently used standardized achievement batteries (Wardrop, 1989). Subtests in CAT-E and F cover both verbal (reading, spelling, and language) and quantitative topics (computation, concepts, and applications). Internal consistency coefficients for virtually all of the subscales exceed .90. Alternate form reliability coefficients are generally in the .80 range.

Missing Data

Six hundred and eighty-four control children were available for assessment in the fall of first grade and self-reports of anxiety were obtained from 570 (83.33%) of these children. Of the 114 children from whom we failed to obtain self-reports of anxiety, there were 34 cases where parents refused to allow their child to participate, 57 cases where parents failed to respond to the consent request in sufficient time to allow for their children to be assessed, and 23 cases where the children were absent from school or otherwise unavailable during the scheduled assessment dates. There were no differences between these 114 children and the 570 children with anxiety data in terms of gender, age, or geographic area.

In the spring of first grade, we were able to locate 531 (77.63%) of the 684 children available in the fall of first grade. Of these 531 children, we were able to obtain self-reports of anxiety from 415 children or 60.67% of the original denominator of 684 children and 72.8% of those children from whom we obtained an anxiety score in fall of first grade. With respect

to the 116 children that were available to us but we were unable to obtain anxiety data from in the spring, there were 12 cases in which parents refused to allow their children to participate, 53 cases where parents did not respond to the consent request, and 51 cases where the children were absent or otherwise unavailable during the allotted assessment dates. We contrasted those children with anxiety scores at both points in time with those with a missing data point in terms of the following baseline characteristics: CAT Math and Reading achievement scores, peer nominations for likability and shy behavior/social participation, and teacher ratings of concentration problems and aggressive and shy behavior/social participation. The only differences found were in terms of CAT Math achievement scores and teacher-rated concentration problems. Children with complete data had significantly higher CAT Math scores than those with missing data [Math, $t(526) = 2.19, p < .05, M = 310.21$ v. $M = 298.10$, range 424, 216 to 640] and significantly less teacher rated concentration problems [concentration problems, $t(591) = -2.70, p < .01, M = 2.56$ v. $M = 2.88$, range 5, 1 to 6]. The differences were relatively trivial, amounting to about $\frac{2}{10}$ of a standard deviation for both math and concentration problems.

RESULTS

Stability of Self-Reported Anxious Symptoms Over Time

As pointed out in the introduction, some have argued that anxious symptoms in young children are relatively transient, and, as such, are insignificant. The following analyses were directed at assessing the stability of anxious symptoms in the study population. A sample of 415 control children was available for retesting on the R-CMAS in the spring, 4 months after the fall testing. The first statistic computed was the intraclass coefficient. The advantage of the intraclass coefficient over the Pearson correlation coefficient in assessing reliability is that it is affected by the magnitude of change over time as well as change in rank (Bartko, 1966). The intraclass correlation coefficient for test–retest performance over the 4-month interval was .64 for boys and .42 for girls, respectively, for the total score. The gender difference between boys and girls was significant ($Z = 3.07, p < .05$). In addition to computing the intraclass coefficient, we also carried out a 2 (Gender) × 2 (Time) repeated-measures analysis of variance (ANOVA) on the total anxiety score, which yielded a significant effect for time [$F(1, 413) = 28.71, p < .0001$]. Specifically, total anxiety scores decreased over the 4-month interval between the initial and second testing (Time 1, $M = 13.41$ $SD = 5.68$; Time 2, $M = 11.69, SD = 6.02$). The magnitude of change was

relatively modest, $\frac{3}{10}$ of a standard deviation. Unlike the intraclass coefficient, the Gender × Time interaction was not significant. A final analysis of test-retest reliability examined stability by rank on the clinical cutoff reported by Reynolds and Richmond (1985).[5] Rank in this analysis refers to whether a child fell below or was at, or above, the cutoff for clinical significance. A Wilcoxon matched-pairs signed-ranks test (Siegel, 1956) was performed to identify the extent to which children changed ranks from the fall to spring assessment. The Z was not significant ($Z = -.7283$, $p = .47$), indicating children's status with respect to the category they fell in (below or at or above the clinical cutoff) was highly stable. Indeed, only 19 out of 415 children changed rank from fall to spring.

Internal Consistency of Self-Reports of Anxious Symptoms Over Time

As noted by Martini, Strayhorn, and Puig-Antich (1990), some have questioned whether prior to the middle elementary school years children are capable of understanding many of the internalizing symptoms included in measures like the R-CMAS. If this were the case, one would expect that first graders' reports of anxious symptoms would demonstrate comparatively low internal consistency, the reason for this hypothesis being that the first graders would be more likely to respond randomly or inconsistently to the items, given that they could not understand the meaning of the items. To test this possibility, coefficient alphas were computed for the entire R-CMAS scale using the fall and spring of first grade data. The alphas for boys were .83 and .83, respectively, and .81 and .82 for girls, respectively. Thus, internal consistency appeared adequate for both boys and girls.

Caseness: The Impact of Anxious Symptoms on Adaptive Functioning over Time

Besides examining whether young children's self-reports of anxious symptoms are internally consistent and stable over time, we also assessed whether they have an impact on adaptive functioning, that is, whether they demonstrate caseness. Separate logistic regressions[6] were performed for

---

[5]Reynolds and Richmond (1985) provided cutoffs for their sample of 6 year olds, separately by race—African—American and white—and gender. A *T*-score of 69, which is approximately 2 *SDs* above the mean, is clinically significant. Since the differences among the subsamples with respect to the raw score equivalent for a *T*-score of 69 were so small, we simply opted to use a single raw score of 27 which is 2 *SDs* above the mean for boys and girls of both races.

[6]The decision to use logistic regression analyses and to dichotomize the variables of interest was based on a number of reasons. First, the distributions of the variables were highly skewed.

each of the indices of adaptive functioning, regressing adaptive functioning at the spring of first grade on self-reports of anxious symptoms from the fall or first grade. Fall levels of adaptive functioning and spring self-reports of anxiety were entered first in the logistic regressions to control for their contribution to spring adaptive functioning. The indices of adaptive functioning included standardized achievement scores, peer nominations of likability and social participation/shy behavior, and teacher-rated concentration problems and aggressive and shy behavior/social participation. Owing to prior evidence of gender differences in the magnitude of relations found between internalizing symptoms and adaptive functioning in children (Lefkowitz and Tesiny, 1985), gender and the Gender × Anxiety interaction were entered into each of the logistic regression analyses. Geographic area was also entered into the logistic models in light of evidence that adaptive functioning may vary as a function of socioeconomic status (SES). For the purpose of the logistic regression analyses, the achievement variable along with the three additional indices of adaptive functioning—teacher-rated concentrated problems and shy and aggressive behavior—and the anxiety scores were converted to dichotomous variables based on a quartile split. The upper quartile was used as the cutpoint for anxious symptoms, peer nominations of shy behavior/social participation, and teacher-rated concentration problems and aggressive behavior. The lower quartile was used as the cutpoint for the achievement construct and peer nominations for likability. With regard to our rationale for the selection of these cutpoints, the children populating the upper quartile of anxious reported a median level of anxious symptoms that fell just short of the number of symptoms associated with Reynolds and Richmond's (1985) suggested cutoff for clinical significance (see footnote 5). The use of a higher cutpoint would have identified too few children for the purpose of these analyses. The majority of children populating the upper quartile of concentration problems in first grade were rated by teachers as in urgent need of educational and/or mental health services. Such was the case for the children rated in the upper quartile of aggression as well. The children failing at or below the bottom

---

Indeed, most children reported few or no anxious symptoms and were rated by their teachers as exhibiting few or no problems with concentration or shy or aggressive behavior. In addition to the highly skewed nature of the distributions, there was severe heteroscedasticity in terms of their joint distributions. Statistical techniques that are based on the assumption of multivariate normality are simply inappropriate with radically nonnormal data such as these. The second reason for our decision to use logistic regression analyses and to dichotomize the variables of interest was based on our interest in examining the impact of a level of anxious symptoms that is likely to require treatment. Third, the normal or average range for most of the indices of adaptive functioning is very wide. Our interest in this study was understanding the degree to which child anxiety heightens the risk for poor achievement and serious problems with attention/concentration, social participation, and aggression in the classroom.

quartile on the achievement variable were in the bottom 20% of the national norms for the CAT.

The results of the logistic regression analyses are presented in terms of the odds ratios.[7] The odds ratio represents the change in the odds of the dependent variable associated with a one-unit change in the independent variable—anxiety—controlling for main effects of geographic area, gender, fall adaptive functioning, and spring anxiety. The relationship between anxious symptoms in the fall and standardized reading and math achievement scores in the spring did reach the conventional level of significance. That is, children in the top quartile of anxiety in the fall of first grade were 7.69 [80% confidence interval (CI) 2.99 to 19.68, $p = .0058$] times more likely to be in the lowest quartile of reading achievement and 2.44 times more likely to be in the lowest quartile in math achievement in the spring of first grade (80% CI 1.49 to 3.98, $p = .02$). Note that whereas the significance level for the relationship between fall anxiety and spring reading achievement was below the Bonferroni correction for the number of regressions performed (.05/7 = .0071429), the relationship between fall anxiety and spring math achievement was not.

Prognostic Value of Adaptive Functioning: Prediction to Later Levels of Anxious Symptoms

We were also interested in understanding the prognostic value of early adaptive functioning with respect to later anxious symptoms. Thus, this next set of analyses was directed at determining the unique contribution of adaptive functioning in the fall of first grade to anxious symptoms in the spring of first grade. Once again logistic regression analyses were utilized and the anxiety and adaptive functioning variables were dichotomized as described above. In a single logistic regression analysis (see footnote 7), spring of first grade self-reported anxiety was regressed on each of the indices of first-grade adaptive functioning in a stepwise fashion, controlling for fall of first grade anxiety and spring of first grade adaptive functioning, gender and geographic area. Of the seven predictors, only fall of first grade teacher-rated aggressive behavior approached significance, falling just short of significance using the Bonferroni correction for alpha

[7]Only the 347 subjects with complete data were used in these analyses. Only relatively trivial differences were found with respect to baseline characteristics between those children with complete data and those missing one or more data points. Children with missing data had significantly higher teacher-rated aggression [$t(548) = -2.97, p = .003, M = 1.78, SD = 1.60$ vs. $M = 1.58, SD = 1.54$, range 5, 1 to 6, .12 SD difference] and concentration problems [$t(548) = -2.73, p = .015, M = 2.84, SD = 1.69$ vs. $M = 2.52, SD = 1.61$, range 5, 1 to 6, .18 SD difference].

(.0071429). Children in the top quartile of aggression in fall of first grade were 2.17 times (80% CI = 1.50 to 3.13, p = .0075) more likely to be in the highest quartile of anxiety in spring of first grade.

Prevalence of Anxious Symptoms as a Function of Geographic Area and Gender

In this last set of analyses, using the data from fall of first grade, we sought to determine the prevalence of anxious symptoms thought to be of clinical significance, based on the cutoff provided by Reynolds and Richmond (1985). That is, we were interested in estimating the number of cases that would likely require intervention or treatment. Logit analysis (Feinberg, 1977), a special class of loglinear analyses in which both the dependent (in this case, anxiety scores above and below the clinical cutoff) and independent variables are categorical, was used to examine whether the prevalence of anxious symptoms varied by geographic area or gender. The interest in variation in rates by geographic area and gender reflected empirical evidence from prior studies (see Angold, 1988). The logit analyses failed to yield either a significant main effect for area or for gender, or for the Area × Gender interaction. The prevalence of clinically significant anxiety was 2.5%.

DISCUSSION

The primary purpose of the present study was to examine the significance of first graders' self-reports of anxious symptoms in terms of their caseness, prevalence, and reliability—internal consistency and stability. As noted in the introduction, stability and caseness are among the criteria outlined by Quay (1986) and others (Beardslee et al., 1985) for judging the validity of a putative psychiatric syndrome, whereas knowledge of the prevalence of a putative disorder provides us with an indication of the extent of the problem and the amount of resources necessary to treat or prevent the problem.

With respect to internal consistency, the coefficient alphas were adequate for both boys and girls, suggesting that children as young as 5 or 6 could understand the meaning of the items on the R-CMAS and answer them reliably. In terms of stability, the intraclass correlation coefficients, repeated-measures ANOVA and Wilcoxon matched-pairs signed-ranks test all provided evidence that anxious symptoms were at least moderately stable over a 4-month interval. Of note, whereas the intraclass coefficient for

boys was relatively strong, it was relatively modest for girls. Thus, when we used a stability statistic which is affected by change in terms of both rank and magnitude, gender differences did appear. Indeed, first-grade anxious symptoms may have greater prognostic value for boys than girls, at least with respect to predicting later levels of anxious symptoms. One possible explanation for this gender difference in stability is that anxious symptoms may have a greater impact on boys' adaptive functioning. Recall that in the introduction we posited that anxious symptoms may impact negatively on children's social and cognitive development, with the resulting deficits in the social and cognitive realms leading to decreases in self-esteem and self-efficacy, which, in turn heighten the risk for subsequent anxious symptoms and disorder. Yet as will be described below we found no gender differences in the impact of anxious symptoms on adaptive functioning. A second possibility is that girls may give off "better signals" that they are distressed than boys, and, therefore, elicit more support from the adult caregiving environment. Indeed, Casey (1993), in a study of children's expression and understanding of emotion, found that girls displayed more positive and negative emotion than boys in response to social feedback.

With respect to the caseness of self-reports of anxious symptoms in first grade, logistic regression analysis was used to determine the unique contribution of fall of first grade anxious symptoms to adaptive functioning the spring of first grade. For both boys and girls, fall anxious symptoms contributed significantly to adaptive functioning in the spring of first grade in terms of reading achievement scores. The relationship between fall anxiety and spring math achievement met the conventional criteria for statistical significance, but not the Bonferroni correction for the number of predictors examined. Besides the failure to find a significant relationship between math achievement and fall anxiety, we failed to find significant relationships between anxious symptoms and peer nominations for likability or shy behavior/social participation. The same was the case for teacher-rated concentration problems, or aggressive and shy behavior/social participation. Given the few statistically significant findings, one might question whether the relationship found between anxious symptoms and reading achievement is a spurious one. However, given the magnitude of the odds ratio and the fact that a Bonferroni correction was used suggests otherwise. Of note, our findings are generally consistent with Kovacs and Goldston's (1991) summary of the literature on the impact of depressive illness on cognitive and social development in children. Indeed, they concluded that most of the reliable evidence of impact was in the area of later academic functioning, with only limited evidence of effects on social or social–cognitive development.

In addition to stability and caseness, the prevalence of anxiety that would likely require intervention or treatment was a third criterion we used to judge the significance of child self-reports of anxious symptoms. It is important to point out, in this regard, that although the rate of a clinically significant anxiety as defined by Reynolds and Richmond (1985) may have been low in our study population, the logistic regression analyses revealed that the impact of anxious symptoms on adaptive functioning was not limited to those children meeting Reynolds and Richmond's cutoff. Indeed, the upper quartile of anxious symptoms was associated with significant impairment in reading achievement. Consequently, the prevalence of anxious symptoms warranting treatment may be greater than what was suggested by the data using Reynolds and Richmond's cutoff.

Besides the issues of caseness, prevalence, and reliability, we were also interested in understanding the contribution of adaptive functioning in the fall to anxious symptoms in the spring. The relationship between teacher ratings of aggressive behavior in the fall and self-reports of anxious symptoms in the spring fell just short of significance when the Bonferroni correction was applied. However, no relationship was found between standardized achievement scores, peer nominations for likability or shy behavior, and spring anxious symptoms. Nor did we find a relationship between teacher-rated concentration problems and shy behavior in the fall and anxious symptoms in the spring. Given that only one predictor even approached significance, it may be the case that the relationship between aggression and anxiety is a spurious one.

Overall, then, first graders' reports of anxious symptoms appear to be at least moderately stable and to have modest to moderate prognostic value with respect to both later anxious symptoms and adaptive functioning. Moreover, the prevalence of clinically significant levels of anxious symptoms—at least in terms of academic achievement—appears to be relatively high. Yet, it is important to note that, although the children lost to followup in the spring of first grade differed only modestly from those who remained in terms of baseline characteristics, they did perform more poorly in terms of academic achievement. Given the relationship found between anxious symptoms and later reading achievement, this raises the possibility that our results may underestimate the caseness of early anxious symptoms. Moreover, the fact that the measure of anxious symptoms employed was a first-stage or screening measure also leads us to believe the degree of caseness found likely represents the lower bound of the true relationship between anxious symptoms and adaptive functioning. A comprehensive psychiatric assessment would likely have allowed not only a more accurate assessment of caseness, but of prevalence and stability as well. Indeed, a highly trained clinician, well versed in interviewing young children, may have been able

to elicit more accurate reports of child symptoms through careful probing of child responses and explanations of the symptoms. The obstacles to mounting such assessments on a large scale, however, are formidable, particularly in terms of cost. The present study was undertaken in an effort to establish whether the costs of such an effort would be justified. Based on our results, we believe the answer should be yes. As Kellam (1990) has noted, a possible approach for future research into the significance of anxious symptoms might be to conduct detailed psychiatric assessments in the context of a multistage sampling design. At the first stage, a relatively economical yet sensitive screening instrument, such as the R-CMAS, could be used to identify cases for more precise and comprehensive assessment at a second or third stage. Ideally, a random stratified sampling procedure would be employed so that representative samples of children from the entire distribution of scores on the first-stage measure of anxious symptoms would be selected for this more precise assessment. Such a design would allow for increasingly precise and comprehensive assessments on smaller but representative subsamples of the larger populations of interest at a reasonable cost.

In conclusion, these data suggest that a reliable and valid assessment of anxiety in children as young as 5 or 6 is quite possible. Moreover, the relationship between anxious symptoms and later reading achievement suggest that such symptoms may warrant mental health services directed at their treatment and prevention. However, as suggested above more comprehensive and precise evaluations of the significance of anxious symptoms and disorders in young children are needed before any firm conclusions can be drawn.

REFERENCES

Angold, A. (1988). Childhood and adolescent depression: I. Epidemiological and aetiological aspects. *British Journal of Psychiatry, 152,* 601-617.

Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin, 83,* 762-765.

Benjamin, R. S., Costello, E. J., & Warren, M. (1990). Anxiety disorders in a pediatric sample. *Journal of Anxiety Disorders, 4,* 293-316.

Beardslee, W., Klerman, G., Keller, M., Lavori, P., & Podorefsky, D. (1985). But are they cases? Validity of DSM-III major depression in children identified in a family study. *American Journal of Psychiatry, 142,* 687-691.

Berkson, J. (1946). Limitations of the application of the fourfold contingency table. *Biometrics Bulletin, 2,* 47-53.

Bird, H., Canino, G., Rubio-Stipec, M., Gould, M. S., Ribera, J., Sesman, M., Woodbury, M., Huertas-Goldman, S., Pagan, A., Sanchez-Lacay, A., & Moscoso, M. (1988). Estimates of prevalence of childhood maladjustment in a community survey in Puerto Rico. *Archives of General Psychiatry, 45,* 1120-1126.

Casey, R. J. (1993). Children's emotional experience: Relations among expression, self-report, and understanding. *Developmental Psychology, 29,* 119-129.

Feinberg, S. E. (1977). *The analysis of cross-classified categorical data.* Cambridge, MA: MIT Press.

Fleming, J., Offord, D., & Boyle, M. H. (1989). The Ontario Child Health Study: Prevalence of childhood and adolescent depression in the community. *British Journal of Psychiatry, 155,* 647-654.

Kellam, S. (1990). Developmental epidemiological framework for family research on depression and aggression. In G. Patterson (Ed.), *Depression and aggression in family interaction* (pp. 11-48). Hillsdale, NJ: Erlbaum.

Kovacs, M. & Goldston, D. (1991). Cognitive and social cognitive development of depressed children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 388-392.

Lapouse, R. & Monk, M. (1959). Fears and worries in a representative sample of children. *American Journal of Orthopsychiatry, 29,* 803-818.

Lefkowitz, M. M., & Tesiny, E. P. (1985). Depression in Children: Prevalence and correlates. *Journal of Consulting and Clinical Psychology, 53,* 647-656.

Martini, D., Strayhorn, J., & Puig-Antich, J. (1990). A symptom self-report measure for preschoolers. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 12-20.

National Institute of Mental Health. (1991). *Implementation of the national plan for research on child and adolescent mental disorders* (Program Announcement PA-91-46). Washington, DC: U.S. Government Printing Office.

Pekarik, E., Prinz, R., Leibert, C., Weintraub, S., and Neale, J. (1976). The pupil evaluation inventory: A sociometric technique for assessing children's social behavior. *Journal of Abnormal Child Psychology, 4,* 83-97.

Quay, H. (1986). Classification. In H. C. Quay & J. S. Werry (Eds., *Psychopathological disorders of childhood* (3rd ed. pp. 1-35). New York: John Wiley & Sons.

Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale (R-CMAS) manual.* Los Angeles: Western Psychological Services.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Wardrop, J. L. (1989). Review of the California Achievement Tests, Forms E and F. In J. Close Conoley & J. Kramer (Eds.), *The tenth mental measurements yearbook* (pp. 128-133), Lincoln: University of Nebraska Press.

Werthamer-Larsson, L., Kellam, S., & Wheeler, L. (1991). Effect of first grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology, 19,* 585-602.