1999

Grade Equivalents: Accuracy and Certainty of Interpretations Among Parents, Teachers, and School Psychologists

Eric R. Smith
Eastern Illinois University

THESIS REPRODUCTION CERTIFICATE

TO: Graduate Degree Candidates (who have written formal theses)

SUBJECT: Permission to Reproduce Theses

The University Library is receiving a number of request from other institutions asking permission to reproduce dissertations for inclusion in their library holdings. Although no copyright laws are involved, we feel that professional courtesy demands that permission be obtained from the author before we allow these to be copied.

PLEASE SIGN ONE OF THE FOLLOWING STATEMENTS:

Booth Library of Eastern Illinois University has my permission to lend my thesis to a reputable college or university or the purpose of copying it for inclusion in that institution's library or research holdings.

[redacted] 7/20/99

Author's Signature Date

I respectfully request Booth Library of Eastern Illinois University **NOT** allow my thesis to be reproduced because:

_____

_____

_____

_____ _____

Author's Signature Date

GRADE EQUIVALENTS: ACCURACY AND CERTAINTY OF INTERPRETATIONS AMONG PARENTS, TEACHERS, AND SCHOOL PSYCHOLOGISTS

BY

ERIC R. SMITH

1974-

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

S.S.P. SPECIALIST IN SCHOOL PSYCHOLOGY

IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY
CHARLESTON, ILLINOIS

1999
YEAR

I HEREBY RECOMMEND THIS THESIS BE ACCEPTED AS FULFILLING THIS PART OF THE GRADUATE DEGREE CITED ABOVE

7/14/99
DATE

[redacted]
ADVISER

7/14/99
DATE

[redacted]
DEPARTMENT HEAD

Abstract

Scores on standardized tests (e.g., Iowa Test of Basic Skills, Wechsler Intelligence Scale for Children III, Wechsler Individual Achievement Test) can be represented by a number of different metrics. One of the ways scores can be represented is with grade equivalents, which tend to be popular with parents and teachers because they seem to be fairly easy to understand. However, several researchers have claimed that grade equivalents are often misinterpreted. Standard scores are viewed by many researchers as the superior type of derived score. However, standardized scores can be difficult to understand, particularly for individuals with little or no training in statistics (e.g., parents, teachers). Thus, grade equivalents are still widely used. Because grade equivalents are used for determining abilities, assessing learning disabilities, and identifying gifted children, it is important to know if they are misinterpreted. The purpose of the current study was to examine the extent to which teachers, school psychologists, and parents misinterpret grade equivalents. The participants included 39 school psychologists, 32 elementary school teachers, and 30 parents. All participants completed a questionnaire that included demographic information (e.g., gender, career) and

several items designed to assess their understanding of the meaning and properties of grade equivalents. Participants were instructed to rank the certainty of their responses on a five-point Likert scale. Although grade equivalents were designed to make standardized test results more meaningful, I found that substantial proportions of parents (76.7%), teachers (67.7%), and even school psychologists (41%) made inaccurate interpretations. In addition, I found that even when people misinterpreted grade equivalents, they generally were as certain of their interpretations as those who accurately interpreted them. Furthermore, in many cases, the people who misinterpreted grade equivalents were certain that they were correct in their interpretations. The results of this study support the notion that grade equivalents are often misinterpreted and should not be used to interpret test scores.

Acknowledgements

I would like to acknowledge Dr. Addison, my thesis advisor, for all the guidance, support, and help he has provided. I would also like to thank Dr. Havey and Dr. Canivez for their assistance with this research project.

In addition, I would like to thank my wife and family for their constant love and support.

Table of Contents

Grade Equivalents: Accuracy and Certainty of Interpretations Among Parents, Teachers, and School Psychologists

Scores on standardized tests (e.g., Iowa Test of Basic Skills, Wechsler Intelligence Scale for Children - Third Edition, Wechsler Individual Achievement Test) can be represented by a number of different metrics. Research has shown that the type of metric used can influence the conclusions that are made (Tindal, Shinn, & Germann, 1987). For example, one of the ways scores can be represented is with grade equivalents, which tend to be popular with parents and teachers because they seem to be fairly easy to understand. However, several researchers (e.g., Anastasi, 1988; Burket, 1984; Hoy & Gregg, 1994; Smith & Addison, 1996) have claimed that grade equivalents are often misinterpreted.

A grade equivalent is the average score that children of a particular grade receive at a specific time in the school year. The months of the school year are represented by a number after a decimal point. For example, if a child obtained a score of 3.8, then that child received the score of the average third grader in the eighth month of the school year (Cohen, Swerdlik, & Smith, 1992).

Grade equivalents are averages rather than standards (Brenner, 1984). Standardized tests are not designed to

reflect whether the skills of a certain grade level have been mastered. For example, the arithmetic section of a standardized test for younger children may consist of problems at a first and second grade level. Therefore, if a child obtained a grade equivalent score a 4.2 on the arithmetic section, it does not mean the child has the math skills of a fourth grader. The score actually means the student performed as well as the average fourth grader in the second month of the school year on a test representing first and second grade math skills (Anastasi, 1988).

In addition, grade equivalents are ordinal data; they are not represented by an interval or ratio scale (Hoy & Gregg, 1994). Therefore, if student A received a grade equivalent of 6.2, and student B received a grade equivalent of 3.1, all that can be appropriately interpreted is that student A scored higher than student B. Because of the properties of ordinal scales, it is not appropriate to suggest that student A scored twice as high as student B. It is also inappropriate to conclude that student A scored three grades higher than student B.

Although the properties of grade equivalents are clear-cut, teachers and parents may not know or understand these properties. In this regard, grade equivalents seem to be susceptible to misinterpretation (Hoy, & Gregg, 1994). For

example, someone who is unaware of the meaning of grade equivalents might see a second grade student's score of 4.0 on arithmetic, and wrongfully conclude that the student performed at a fourth grade level. The tendency for people to make this conclusion is supported by the results obtained by Smith and Addison (1996). They examined whether grade equivalents were misinterpreted by graduate students in school psychology, elementary education majors, and students in a general education course. They found that the grade equivalents were misinterpreted by 53.7% of the graduate students, 87.9% of the elementary education majors, and 68.2% of the students in a general education course.

An additional problem with grade equivalent scores is the fact that they do not accurately measure growth in academic achievement. Research has shown that the rate of learning between grades is irregular (Beggs & Hieronymus, 1968); however, grade equivalents assume that learning occurs in a constant manner (Berk, 1981; Reynolds, 1981). Another underlying assumption of grade equivalents is that growth only occurs 10 months out of the year (Burket, 1984). If this were true, no growth or loss would occur over the summer. However, research has shown that because of forgetting, achievement loss tends to occur over the summer months (Beggs & Hieronymus, 1968).

One of the reasons grade equivalents can not be used to measure growth accurately is related to the way they are derived. In order for learning to remain constant throughout the school year, the lower and upper ends of the scale are extrapolated (Reynolds, 1981). Extrapolation distributes the grade equivalents in a manner which reflects that learning occurs in a linear fashion. However, research has indicated that learning is a process that occurs sporadically, with sudden growth spurts (Beggs & Hieronymus, 1968).

In addition, because tests are not administered during every month of the school year, scores must be interpolated to complete the gaps between administrations (Reynolds, 1981). This estimation process produces scores that are generally too low in the fall, and scores that are too high in the spring. Therefore, if academic screening is done in the fall there is a greater chance of falsely identifying children as learning disabled. Conversely, if screening is done in the spring there is a greater chance of failing to identify learning disabilities that exist (Berk, 1981).

Another problem with grade equivalents is that they tend to exaggerate small differences in performance (Angoff, 1971). Grade equivalent scores that are slightly above or below the mean can show up as one or more grade levels above

or below the child's grade level (Berk, 1981). For example, if a sixth grader is tested in the seventh month of the school year and receives a standard score of 90 (M = 100, SD = 15) on the Oral Expression subtest of the Wechsler Individual Achievement Test, that student would receive a grade equivalent of 3.7. Even though the performance falls within the average range, the grade equivalent is three grades below placement.

An alternative to grade equivalents for interpreting standardized test scores is percentile ranks. Percentile ranks have several advantages when compared to grade equivalents or standardized scores. They are fairly easy to understand and they are useful in making relative comparisons (Lien, 1976). In addition, percentile ranks can be used for a wide variety of tests, such as tests of ability, personality, intelligence, etc. (Cohen et al., 1992).

However, there are several shortcomings associated with the use of percentile ranks. One disadvantage is that percentile ranks are also ordinal data. Therefore, no arithmetic operations can be performed on the data. Furthermore, if the distribution of scores approximate a normal curve, as in most standardized tests, then the differences between raw scores are exaggerated near the

median of the distribution and minimized near the extreme ends of the distribution (Gronlund, 1976).

In addition, research has shown that percentile ranks have been misinterpreted by school psychologists (Huebner, 1989; Ross, 1990). Huebner found that if the test score was at 1 standard deviation below the mean, school psychologists were more likely to falsely identify learning disabilities when percentiles were used rather than grade equivalents or deviation IQs. Furthermore, research has shown that school psychologists tend to view percentiles as interval rather than ordinal data (Huebner, 1989; Ross, 1990).

Another alternative to grade equivalents is to use standard scores. The use of standard scores for interpreting the results on standardized tests has many advantages. The main advantage is that standard scores represent an interval scale; therefore, arithmetic operations can be performed on them. These operations can be particularly useful in discrepancy analysis (Reynolds, 1981). In addition, standard scores have the same percentile rank across age because they are based on the variability in scores about the mean at each age level. For example, a score that falls one standard deviation below the mean has a percentile rank of 16 at every age. Conversely, a score falling one grade level below the average grade

level has a different percentile rank at every age (Reynolds, 1981). Additionally, standard scores are more accurate and precise than grade equivalents. Unlike the case with grade equivalents, it is usually not necessary to interpolate scores to arrive at exact score points during the construction of tables for the conversion of raw scores to standard scores. Another advantage of standard scores is that extrapolation is usually unnecessary for 99% of the scores (i.e., scores within three standard deviations of the mean) (Reynolds, 1981).

Standard scores are viewed by many researchers as the superior type of derived score (Anastasi, 1988; Burket, 1984; Berk, 1981; Reynolds, 1981). However, standardized scores can be difficult to understand, particularly for individuals with little or no training in statistics (e.g., parents, teachers). Thus, grade equivalents are still widely used.

Because grade equivalents are used for determining abilities, assessing learning disabilities, and identifying gifted children, it is important to know if they are misinterpreted. In a previous study it was demonstrated that grade equivalents were misinterpreted by graduate students in school psychology, elementary education majors, and students in a general education course (Smith & Addison,

1996). The purpose of the current study was to examine the extent to which teachers, school psychologists, and parents misinterpret grade equivalents.

Because teachers and school psychologists are likely to have received some training in interpreting grade equivalents, I expected them to interpret grade equivalents more accurately and be more certain of their responses than the parents. Additionally, school psychologists were expected to interpret grade equivalents more accurately and be more certain of their responses than both teachers and parents because of school psychologists' training and experience with interpreting test scores. Additionally, this study was designed to examine whether individuals who accurately interpret grade equivalents are more certain of their interpretations than those who misinterpret them. I expected the individuals who accurately interpreted grade equivalents to be more certain of their interpretations than those who misinterpreted them.

Method

Participants

The participants included 39 school psychologists, 32 elementary school teachers, and 30 parents. Thirty-two (82.1%) of the school psychologists were women. The mean age reported by the school psychologists was 38.97 (*SD* =

10.65) and the average years of experience was 10.44 (*SD* = 8.64). Twenty-eight (87.5%) of the teachers were women. The mean age reported by the teachers was 39.21 (*SD* = 10.62) and the average years of experience was 15.03 (*SD* = 9.58). Twenty-six (86.7%) of the parents were women. The mean age reported by the parents was 36.58 (*SD* = 7.45).

Materials

All participants completed a questionnaire that included demographic information (e.g., gender, career) and several items designed to assess their understanding of the meaning and properties of grade equivalents. Participants were instructed to rank the certainty of their responses on a five-point Likert scale (1=extremely uncertain, 2=somewhat uncertain, 3=neutral, 4=somewhat certain, and 5=extremely certain) (see the Appendix for a copy of the questionnaire).

Each item consisted of a brief scenario followed by five possible interpretations. Instructions were as follows: "Please circle from the following all of the accurate interpretations (you may circle more than one) and indicate the certainty of your response below." There were two correct interpretations and three incorrect interpretations for each item on the questionnaire. If the participant chose at least one of the correct interpretations and did not choose any of the incorrect

interpretations, the item was scored as correct. If the participant chose one or more of the incorrect interpretations, the item was scored as incorrect regardless of how many correct interpretations were chosen.

Procedure

The interns of the school psychology graduate program at Eastern Illinois University were supplied with questionnaires and self-addressed stamped envelopes. They distributed the questionnaires and envelopes to their supervisors and other school psychologists at their internship sites.

In order to obtain teachers and parents for the study, the principals of several elementary schools were contacted and briefly explained the study. If permission was granted, the questionnaires were given to the secretaries at the various schools with instructions to disseminate them to the teachers. Completed questionnaires were collected at a later time. At one school, the principal gave me permission to have some of their teachers send questionnaires to the parents of their students. Those teachers then returned the questionnaires to the school's secretary after they were completed.

To ensure the anonymity of the participants, the questionnaires did not contain items that could be used to

identify them. The overall return rate was 48.3%.

Results

Chi square tests of independence were used to examine the relationship between accuracy of response and group status (i.e., school psychologists, teachers, and parents) for each scenario (see Table 1). An alpha level of .05 was used for all of the statistical analyses. Significant differences among the groups occurred on all three scenarios (Scenario one, $\chi^2(2, N=100) = 11.46$, $p = .0033$; scenario two, $\chi^2(2, N = 100) = 9.51$, $p = .0086$); scenario three, $\chi^2(2, N=100) = 11.96$, $p = .0025$.

Descriptive statistics shown in Table 2 show that for scenario one, 33.3% of the school psychologists, 68.8% of the teachers, and 66.7% of the parents interpreted the grade equivalent inaccurately. For scenario two, 51.3% of the school psychologists, 65.6% of the teachers, and 86.7% of the parents interpreted the grade equivalent inaccurately. For scenario three, 38.5% of the school psychologists, 68.8% of the teachers, and 76.7% of the parents interpreted the grade equivalent inaccurately. Across all three scenarios, the percentage of inaccurate interpretations was 41% for the school psychologists, 67.7% for the teachers, and 76.7% for the parents.

For each scenario, each participant was placed into one

of two categories: those who interpreted the scenario accurately, and those who interpreted the scenario inaccurately. A 2(accurate vs. inaccurate interpretation) X 3(school psychologists vs. teachers vs. parents) analysis of variance was conducted for each scenario to examine the interaction effect of accuracy of interpretation and group status on certainty of response (see Tables 3 and 4). This analysis was used to determine whether participants who accurately interpreted grade equivalents were more certain of their responses than the participants who misinterpreted grade equivalents, which groups were more certain of their responses, and whether there was an interaction effect of accuracy and group status on certainty of response. An alpha level of .05 was used for all comparisons.

For scenario one, the main effect of group status on certainty of response was significant, $F(2, 100) = 7.07$, $p = .0014$. Scheffe's Multiple Comparisons were used to examine the differences in certainty of response ratings among the three groups. The only significant difference was between the school psychologists ($M = 4.36$, $SD = .84$) and the parents ($M = 3.50$, $SD = 1.14$) $p = .002$. The main effect of accuracy of response on certainty and the interaction were not significant.

On scenario two, the main effect of group status on

certainty of response was significant, $F(2, 100) = 7.00$, $p = .0015$. Scheffe's Multiple Comparisons were use to examine the differences in certainty of response ratings among the three groups. Again, the only significant difference was between the school psychologists ($M = 4.31$, $SD = .89$) and the parents ($M = 3.47$, $SD = 1.07$) $p = .002$. The main effect of accuracy of response on certainty of response was significant, $F(1, 100) = 6.40$, $p = .0079$. Participants who interpreted scenario two accurately ($M = 4.32$, $SD = .68$) were significantly more certain than the participants who interpreted scenario two inaccurately ($M = 3.79$, $SD = 1.11$). The interaction was not significant.

On scenario three, the main effect of group status on certainty of response was significant, $F(2, 100) = 7.67$, $p = .0002$. Scheffe's Multiple Comparisons were use to examine the differences in certainty of response ratings between the three groups. Consistent with the results for the other scenarios, the school psychologists ($M = 4.38$, $SD = .71$) were significantly more certain of their interpretations than the parents ($M = 3.43$, $SD = 1.07$) $p < .001$. However, the interaction was significant, $F(2, 100) = 3.37$, $p = .0386$. Thus, on scenario three, the effect of accuracy of response on certainty ratings depended on the participant group. Teachers and school psychologists who interpreted

scenario three accurately were significantly more certain of their responses than parents who interpreted scenario three accurately. Additionally, parents, teachers, and school psychologists who interpreted scenario three inaccurately were also significantly more certain of their responses than parents who interpreted scenario three accurately (see Figure 1). The main effect of accuracy of response on certainty ratings was not significant.

Discussion

Because teachers and school psychologists are likely to have received some training on interpreting grade equivalents, it was predicted that they would interpret grade equivalents more accurately than the parents. However, the results show that the teachers interpreted grade equivalents more accurately than did the parents on scenario two only. Thirty-four percent of the teachers and 13% of the parents interpreted this scenario accurately. On scenario two, the participants were asked to compare two grade equivalent scores from two separate students. This item was designed to examine the participants' understanding of the ordinal properties of grade equivalents and their awareness that grade equivalents are averages rather than standards. Thus, one explanation for the difference is that teachers are more aware of the interpretive limitations for making comparisons than parents.

I predicted that the school psychologists would interpret grade equivalents more accurately than both parents and teachers because of school psychologists' extensive training and experience with interpreting test scores. The results were consistent with this prediction in that school psychologists interpreted grade equivalents more accurately than did parents on all three scenarios. Additionally, school psychologists interpreted grade equivalents more accurately than teachers on scenarios one and three. On scenario one and three, the participants were asked to interpret the grade equivalent scores from an individual. On scenario two, the participants were asked to compare grade equivalents from two students. Therefore, school psychologists were generally more accurate at interpreting grade equivalents than were teachers; however, there was no significant difference in their accuracy of interpretation when they were asked to compare grade equivalents from two students, they were as accurate as teachers.

This study also examined the certainty of the participants' answers by having them rank their certainty on a Likert scale. I expected the school psychologists to be more certain of their responses than both parents and teachers because their extensive training and experience with interpreting test scores. However, school

psychologists were more certain of their responses only when compared to parents. Because teachers are likely to have some training and experience with interpreting grade equivalents, I expected them to be more certain of their responses than the parents. However, the teachers were not more certain of their responses than were parents.

In addition, this study examined whether individuals who accurately interpreted grade equivalents were more certain of their interpretations than the individuals who misinterpreted them. I expected that participants who made accurate interpretations would be more certain of their responses than participants who made inaccurate interpretations. This hypothesis was supported only for scenario two. Therefore, participants who accurately compared two students were more certain of their responses than those who inaccurately compared the students. A possible explanation for this finding is that parents, teachers, and school psychologists all recognize whether or not they possess the knowledge to accurately use grade equivalents to compare two students.

On scenarios one and three, participants who accurately interpreted grade equivalents were no more certain of their responses than participants who misinterpreted them. One explanation for this finding is that grade equivalents are so easily misinterpreted that even when individuals

misinterpret them those individuals are certain their interpretations are accurate.

Nevertheless, there was an interaction effect for scenario three. The effect of accuracy of response on certainty ratings depended on the participant group. Teachers and school psychologists who interpreted the scenario accurately were significantly more certain of their responses than parents who interpreted it accurately (see Figure 1).

Grade equivalents were designed to make standardized test results more meaningful. However, substantial proportions of parents (76.7%), teachers (67.7%), and even school psychologists (41%) made inaccurate interpretations. In addition, even when people misinterpreted grade equivalents, they generally were as certain of their interpretations as those who accurately interpreted them. Furthermore, in many cases, individuals who misinterpreted grade equivalents were certain that they were correct in their interpretations. These results support the notion that grade equivalents are often misinterpreted and should not be used to interpret test scores. Another implication of my study is that universities need to stress the meaning of grade equivalents to education majors and to graduate students in school psychology programs. These conclusions are consistent with those from a previous study of

undergraduate and graduate students in education and school psychology (Smith & Addison, 1996).

In conclusion, grade equivalents are often misinterpreted and probably should not be used. The best alternative may be to use standard scores. Although standard scores are viewed by many researchers as the superior type of derived score for interpreting standardized test results (e.g., Anastasi, 1988; Berk, 1981; Burket, 1984; Reynolds, 1981), research should be done to determine whether they can be presented in a meaningful manner to parents and teachers.

References

Anastasi, A. (1988). Psychological and testing (6th ed.). New York: Macmillan.

Beggs, D. L., & Hieronymus, A. N. (1968). Uniformity of growth in the basic skills through the school year and during the summer. Journal of Educational Measurement, 5, 91-97.

Berk, R. A. (1981). What's wrong with using grade equivalent scores to identify LD children. Academic Therapy, 7(2), 133-140.

Brenner, K. (1984). The psychological impact of grade norms on children. Psychology: A Quarterly Journal of Human Behavior, 21(1), 50-51.

Burket, G. (1984). Response to Hoover. Educational Measurement: Issues and Practice, 3(4), 15-16.

Cohen, R. J., Swerdlik, M. E., & Smith, D. K. (1992). Psychological Testing and Assessment (2nd ed.). Mountain View, CA: Mayfield.

Gronlund, N. E. (1965). Measurement and evaluation in teaching. New York: Macmillan.

Hoy, C., & Gregg, N. (1994). Assessment: The special educator's role. Belmont, CA: Brooks/Cole.

Huebner, S. (1989). Errors in decision making: a comparison of school psychologists interpretations of grade equivalents, percentiles, and deviation IQs. School

Psychology Review, 18(1), 51-55.

Lien, A. J. (1976). Measurement and evaluation of learning (3rd ed.). Dubuque, IA: Wm. C. Brown.

Reynolds, C. R. (1981). The fallacy of "Two years below the grade level for age" as a diagnostic criterion for reading disorders. Journal of School Psychology, 19(4), 350-358.

Ross, R. P. (1990). Consistency among school psychologists in evaluating discrepancy scores: A preliminary study. Learning Disability Quarterly, 13, 209-219.

Seltzer, M. H., Frank, K. A., & Bryk, A. S. (1994). The metric matters: The sensitivity of conclusions about growth in student achievement to choice of metric. Educational Evaluation and Policy Analysis, 16(1), 41-49.

Smith, E. R., & Addison, W. E. (1996). Grade equivalents: Are they misinterpreted? Paper presented at the annual meeting of the Illinois School Psychologists Association, Peoria, IL.

Tindal, G., Shinn, M., & Germann, G. (1987). The effect of different metrics on interpretation of change in program evaluations. Remedial and Special Education, 8(5), 19-28.

Appendix

Grade Equivalent Questionnaire

1. Gender: **M** **F**

2. Age: _____

3. Career: _____ For how long: _____

4. County: _____

Instructions: Imagine you are asked to interpret the results on a standardized test and the scores of the test are reported using grade equivalents, a statistic used for interpreting scores. This questionnaire contains four brief scenarios. After each scenario are five possible interpretations. Circle all of the accurate interpretations (you may circle more than one) and indicate the certainty of your response below. Please answer all four items to the best of your ability, even if you are unsure of your answer. After you answer each item, indicate the certainty of your answer. The results of your survey are fully confidential. Your participation is greatly appreciated.

1. On a standardized test, Jaimie, a child in the fourth grade receives a grade equivalent score of 6.2 on mathematic ability. Please circle from the following all of the accurate interpretations (you may circle more than one) and indicate the certainty of your response below.

a. Jaimie is performing at a sixth grade second month level in mathematics.

b. Jaimie did as well as the average sixth grader in the second month of the school year.

c. Jaimie has mastered 4th and 5th grade math skills.

d. Jaimie needs to be challenged with more difficult math problems.

e. Jaimie is above average but may be at a fourth grade level in mathematics.

Please indicate the certainty of your answer; circle one.

extremely uncertain	somewhat uncertain	neutral	somewhat certain	extremely certain

2. On the reading ability section of the Iowa Test of Basic Skills, two third grade students receive grade equivalent scores. Johnny receives a score of 2.1 and James receives a score of 4.2. Please circle from the following all of the accurate interpretations (you may circle more than one) and indicate the certainty of your response below.

a. James did better than Johnny

b. Johnny did as well as the average second grader in the first month of the school year and James did as well as the average 4th grader in the second month of the school year.

c. Johnny performed at a second grade level and James performed at a fourth grade level.

d. James did twice as well as Johnny.

e. James scored two grades higher than Johnny.

Please indicate the certainty of your answer; circle one.

extremely uncertain	somewhat uncertain	neutral	somewhat certain	extremely certain

3. On a standardized test Kathy, a sixth grader, receives a grade equivalent score of 4.2 on mathematic ability. Please circle from the following all of the accurate interpretations (you may circle more than one) and indicate the certainty of your response below.

a. Kathy has difficulty with mathematics

b. This score provides evidence that Kathy needs special education in the area of mathematics

c. Kathy is at a fourth grade level in mathematics

d. Kathy performed as well as the average fourth grader in the second month of the school year

e. Kathy is below average but may be able to solve sixth grade mathematics problems without any great difficulty.

Please indicate the certainty of your answer; circle one.

extremely uncertain	somewhat uncertain	neutral	somewhat certain	extremely certain

Thanks for your participation.

Table 1

Analysis of Accuracy by Group

Scenario	n	df	
1.	101	2	
Chi Square			11.4556
Probability			.0033
2.	101	2	
Chi Square			9.5179
Probability			.0086
3.	101	2	
Chi Square			11.9592
Probability			.0025

Table 2

Percentages of Incorrect and Correct Responses by Group

Group	Incorrect	Correct
Scenario		
1		
Parents (n=30)		
Observed Frequency	20	10
Percentage	66.7%	33.3%
Teachers (n=32)		
Observed Frequency	22	10
Percentage	68.8%	31.2%
School Psychologists (n=39)		
Observed Frequency	13	26
Percentage	33.3%	66.7%
2		
Parents (n=30)		
Observed Frequency	26	4
Percentage	86.7%	13.3%
Teachers (n=32)		
Observed Frequency	21	11
Percentage	65.6%	34.4%
School Psychologists (n=39)		
Observed Frequency	20	19
Percentage	51.3%	48.7%
3		
Parents (n=30)		
Observed Frequency	23	7
Percentage	76.7	23.3%
Teachers (n=32)		
Observed Frequency	22	10
Percentage	68.8%	31.2%
School Psychologists (n=39)		
Observed Frequency	15	24
Percentage	38.5%	61.5%

Table 3

Mean Scores for Certainty of Response by Group

Scenario		Group		
		Parents (n=30)	Teachers (n=32)	School Psychologists (n=39)
1.				
M		3.50	3.94	4.36
SD		1.14	.88	.84
F-Ratio	7.07			
Probability	.0014			
2.				
M		3.47	4.03	4.31
SD		1.07	.93	.89
F-Ratio	7.00			
Probability	.0015			
3.				
M		3.43	3.97	4.38
SD		1.07	1.00	.71
F-Ratio	9.39			
Probability	.0002			

Table 4

Mean Scores for Certainty Ratings Based on Accuracy

Scenario		Accuracy: Yes	Accuracy: No
1. (n=101)			
M		4.13	3.83
SD		.88	1.09
F-Ratio	2.45		
Probability	.1208		
2. (n=101)			
M		4.32	3.79
SD		.68	1.11
F-Ratio	7.35		
Probability	.0079		
3. (n=101)			
M		4.05	3.92
SD		1.01	.99
F-Ratio	.53		
Probability	.4701		

Table 5

Incorrect Interpretations Endorsed by Participants who Misinterpreted Grade Equivalents

Group	Scenario	Incorrect Interpretations		
	1	(A)	(C)	(D)
Parents (n=20)				
Observed Frequency		16	9	6
Percentage		80%	45%	30%
Teachers (n=22)				
Observed Frequency		16	9	11
Percentage		72.7%	40.9%	50%
School Psychologists (n=13)				
Observed Frequency		7	3	6
Percentage		53.8%	23.1%	46.2%
	2	(C)	(D)	(E)
Parents (n=26)				
Observed Frequency		17	12	15
Percentage		65.4%	46.2%	57.7%
Teachers (n=21)				
Observed Frequency		14	4	13
Percentage		66.7%	19%	61.9%
School Psychologists (n=20)				
Observed Frequency		9	0	14
Percentage		45%	0%	7%
	3	(A)	(B)	(D)
Parents (n=23)				
Observed Frequency		13	9	22
Percentage		56.5%	39.1%	95.7%
Teachers (n=22)				
Observed Frequency		17	0	11
Percentage		77.3%	0%	50%
School Psychologists (n=15)				
Observed Frequency		10	0	7
Percentage		66.7%	0%	46.7%

Figure 1

Interaction Effect on Scenario 3