Journal of
Education and Human Development ISSN 1934-7200
Volume 2, Issue 2, 2008

The WISC-IV General Ability Index in a Non-clinical Sample

Gail M. Cheramie, Associate Professor, University of Houston-Clear Lake, cheramie@uhcl.edu
Mary E. Stafford, Associate Professor, University of Houston-Clear Lake, stafford@uhcl.edu
Sarah S. Mire, Licensed Specialist in School Psychology, Graduate Student – University of Houston, sarahmire@gmail.com

Abstract

The Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) yields a Full Scale IQ (FSIQ) based on ten subtests. An alternative global score, the General Ability Index (GAI), can be calculated when the profile meets certain criteria that involve variability with respect to the cognitive factors of working memory and processing speed. The prevalence of the GAI in clinical samples of children (e.g., students with learning disabilities and attention disorders) has been established. The purpose of this study was to investigate the prevalence of children who met the GAI criteria in a non-clinical sample and to determine which of the criteria used for the GAI are most frequently met. The results indicate the prevalence of meeting the GAI criteria is high in a non-clinical sample and similar to results obtained in clinical samples, most non-clinical students have a GAI score higher than the FSIQ. These results imply that the GAI cannot be associated only with clinical samples.

Introduction

The Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) is one of the most widely used measures for the assessment of cognitive abilities in school-age children and adolescents. The WISC-IV consists of ten core subtests that combine to form four factors or indices – Verbal Comprehension Index (VCI; Similarities, Vocabulary, Comprehension), Perceptual Reasoning Index (PRI; Block Design, Picture Concepts, Matrix Reasoning), Working Memory Index (WMI; Digit Span, Letter-Number Sequencing), and Processing Speed Index (PSI; Coding, Symbol Search). These ten subtests also result in the calculation of a global composite measure, the Full Scale Intelligence Quotient (FSIQ). The General Ability Index (GAI) was developed as an alternative global measure of cognitive functioning designed to reduce the effects of working memory and processing speed (Raiford, Weiss, Rolfhus, & Coalson, 2005), as these factors have been shown to be lower in samples of students with disabilities thus having the potential to attenuate the FSIQ. Since several states include an aptitude-achievement discrepancy model for determination of learning disabilities based on the FSIQ compared to achievement scores, a lowered FSIQ in such samples can have an impact on classification. The GAI is calculated by summing the six subtests that comprise the VCI and PRI.

The GAI was developed for the WISC-III by Prifitera, Weiss, and Saklofske (1998). According to Prifitera, et. al. (1998), research had shown that certain clinical groups (e.g., students with learning disabilities, students with attention deficit hyperactivity disorders) were likely to score lower on the Arithmetic and Coding subtests, which were used in the calculation of the WISC-III FSIQ. To reduce the effects of these subtests on the FSIQ, the GAI was calculated by adding up the remaining eight subtests. With the revision of the WISC-III, the Verbal and Performance IQ scales were eliminated and replaced with the VCI and PRI. In addition, four subtests (Digit Span, Letter-Number Sequencing, Coding, and Symbol Search) as compared to two (Arithmetic and Coding) are involved in the cognitive areas of working memory and processing speed and are used in the calculation of the WISC-IV FSIQ. Thus, the WISC-IV FSIQ is more affected by these cognitive processes. The GAI was developed to provide a summary score that is less sensitive to the effects of working memory and processing speed.

Studies performed during the WISC-IV standardization process (as reported in Raiford, et. al., 2005) indicated that a larger percentage of children in certain clinical groups (Reading Disorder, Reading and Written Expression disorders, Reading, Written Expression, and Mathematics disorders, Learning Disorder, and Attention-Deficit/Hyperactivity Disorder) had a FSIQ less than the GAI. The GAI could, therefore, be a more appropriate aggregate for these groups since the processes associated with the disability condition may have attenuated the FSIQ. The calculation of the GAI is not restricted, however, to use for students in certain clinical groups. Basically, the GAI is used in situations where there is substantial variability within the cognitive profile. In addition, the GAI can be used interpretively to show the effects of working memory and processing speed on the full expression of cognitive ability.

In the WISC-IV Technical Report #4 (Raiford, et. al, 2005), criteria were provided for when the GAI should be considered in clinical situations. These rules are as follows:

1. When there is a significant and unusual difference between the VCI and WMI,
2. When there is a significant and unusual difference between the PRI and PSI,
3. When there is a significant and unusual difference between the WMI and PSI, or
4. When there is significant and unusual variability between the subtests that comprise the WMI or PSI.

A statistically significant difference between index scores is determined by using the pair-wise discrepancy comparisons provided in Table B.1 in the WISC-IV Administration and Scoring Manual (Wechsler, 2003, p. 256). Table B.2 (Wechsler, 2003, p. 257) provides the data on base rates for the pair-wise index score discrepancies. Tables B.3 and B.4 provide the data to determine significance and base rates for subtest comparisons (Wechsler, 2003, pp. 263, 264). As noted above, the GAI is considered when there is a "significant" and "unusual" difference between index scores, or between the subtests which comprise the WMI and PSI. A significant difference is determined by the critical values at the .15 and .05 levels, while an unusual difference is determined by the base rate. The term unusual is operationally defined as occurring in less than 10 to 15% of the standardization sample (Sattler, 2001).

While the GAI can be beneficial for interpretive purposes, can be used in educational decisions, and is prevalent in clinical samples, what is the prevalence of the GAI profile in non-clinical samples? This was the guiding question for this investigation. The following questions were developed for the study: (a) What percentage of students in a non-referred, non-clinical sample would meet the criteria for the GAI?; (b) Which of the four criteria used would most frequently be met?; and (c) What effect does using a less than 10 or less than 15% base rate have on the number of cases that meet the criteria for the GAI?

Method

Participants

The WISC-IV was administered by graduate students in an intellectual assessment course and a database was created from the protocols. As part of the course requirement, no child assessed could be receiving special education services or have a clinical diagnosis. A total of 461 WISC-IV protocols comprised the sample for this investigation. The children tested ranged in age from 6-0 to 16-11, and the number of children per age level was relatively evenly distributed across the age range (Age 6: n=45; Age 7: n=35; Age 8: n=42; Age 9: n=31; Age 10: n=41; Age 11: n=53; Age 12: n=48; Age 13: n=42; Age 14: n=28; Age 15: n=42; Age 16: n=54). There were 235 female and 190 male children. Gender was not available for 36 children. [A limitation of this study is the absence of data regarding the ethnic status of the children who were evaluated.]

Instrument

The WISC-IV was the instrument used in this investigation. All ten core subtests had to have been administered (no substitutions) and were used to calculate the VCI, PRI, WMI, PSI, and FSIQ.

Procedures

Protocols were selected from archival data thus minimizing any bias in protocol scoring and selection. As part of the intellectual assessment course, each protocol is checked for clerical and mathematical errors and graded by the course professor. In addition, each graduate student is directly observed in the standardized administration of the WISC-IV. Only the protocols of those students who administered and scored the test with criterion levels of accuracy (90% and above) were selected for the study.

The GAI was calculated by summing the scaled scores on the Similarities, Vocabulary, Comprehension, Block Design, Picture Concepts, and Matrix Reasoning subtests; these subtests comprise the VCI and PRI. A GAI was calculated for each protocol in order to investigate the overall correlation of the GAI with the FSIQ. The table for the GAI score conversion is included in the WISC-IV Technical Report #4 (Raiford et. al., 2005, p. 5).

Protocols were evaluated to determine if they met any of the four criteria identified above (VCI-WMI, PRI-PSI, WMI-PSI, subtest variability within the WMI or PSI). The criterion met had to be statistically significant and unusual at either the 15% or 10% base rate. The level of statistical significance used was .05 (Tables B.1 and B.3, Wechsler, 2003). If a protocol met any of the criteria at the 10% base rate, then that protocol was designated as meeting the criterion at the 15% base rate as well. Based on Tables B.2 and B.4, the following critical values were used to determine base rates at the 10% and 15% levels: VCI-WMI: <10%=19 points and <15%=15 points; PRI-PSI: <10%=20 points and <15% = 16 points; WMI-PSI: <10%=21 or 22 points and <15% =18 points; and subtest variability within the WMI and PSI=4 points.

Results

For the three questions that were addressed in this study (what percentage of students in a non-referred, non-clinical sample met the rules for the GAI; which of the four rules used were met; what effect does using a less than 10 or less than 15% criterion have on the number of students that meet the rules for the GAI), results are given in the paragraphs below, respectively.

Percentage of Non-Clinical Students Meeting the Rules for GAI

In this sample of 461 children, 73.1% (337 children) met the rules for use of the GAI. An analysis of the difference between the GAI and FSIQ for those meeting the GAI rules revealed that in 62% of the protocols the GAI was higher than the FSIQ, whereas in 32.7% of the protocols the GAI was lower than the FSIQ. Thus, the majority of this non-clinical sample had a FSIQ lower than the GAI. The percentage of protocols where the GAI was higher than the FSIQ (62%) is roughly equivalent to the corresponding percentage in the clinical sample (70%; Raiford et al., 2005, p. 3).

Which Rules Were Met

Table 1 contains the four rules for using the GAI, as well as combinations of rules, and the sample sizes of children who met each rule/group of rules. It should be noted that those children meeting each rule include both children who met only the specific rule (i.e., 38 children met the rule for a significant difference between the VCI and the WMI, Rule #1), as well those who met a combination of rules that included rule #1. The investigators were interested in the relative frequency with which children met each of the rules for using the GAI. In the next paragraphs, the number of children who met each rule will be reported as (a) those who met only the rule and (b) those who met the rule and at least one other rule.

Of the 337 children who met at least one rule for using the GAI, the largest number of children (181) met rule #2 (i.e., PRI-PSI discrepancy). Of those 181 children, 50 met rule #2 alone, whereas 131 met both rule # 2 and at least one other rule. Rule #4 (i.e., scatter within the WMI/PSI subtests) produced the second largest number (150), with 51 children meeting only rule #4 and 99 children meeting rule #4 as well as at least one other rule.

Rules # 1 (i.e., VCI-WMI discrepancy) and #3 (i.e., WMI-PSI discrepancy) produced approximately equal number of instances (130 and 132, respectively). Thirty-eight children met rule #1 alone; whereas 92 met rule #1 and at least one other rule. Sixteen children met Rule #3 alone, whereas 116 met rule #3 and at least one other rule.

Table 1. Frequency of Children Meeting each Rule/Rules.

Rule	**# of Children**	**Multiple Rules**	**# of Children**
1. VCI – WMI	38	Rules 2 and 3	32
2. PRI – PSI	50	Rules 2 and 4	25
3. WMI – PSI	16	Rules 3 and 4	8
4. Scatter within WMI/PSI	51	Rules 1, 2, 3 and 4	8
		Rules 1, 2, and 3	17
Multiple Rules		Rules 1, 2, and 4	5
Rules 1 and 2	19	Rules 2, 3, and 4	25
Rules 1 and 3	15	Rules 1, 3, and 4	11
Rules 1 and 4	17	Total	337

Effect of Base Rate Level on Number of Children Meeting Rules for GAI

Table 2 contains data related to the base rate level of less than 10% or less than 15% on meeting the rules for using the GAI. Base rate data from the standardization sample were calculated for each rule, using the less than 10% level as well as using the less than 15% level. Frequencies reflect overall numbers of children, regardless of type criteria that was met (as found in Table 1). Results indicate that 73.1% of the sample (337 children) fell within the range of meeting the rules for using the GAI. Of this total percentage, 76% (255 children) met the criterion of less than 10%, whereas 24% (82) fell within the 10% to 14.99% range.

Table 2. Frequency of Base Rate Levels of Less than 10% and Less than 15%

Met Criterion at	**Frequency**	**Percentage**
0 to 9.99%	255	76%
10% to 14.99%	82	24%
Total	337	100%

Discussion

Although the GAI has been traditionally associated with clinical samples, the results of this study indicate that the prevalence of children who meet the GAI criteria in non-clinical samples is also high. The prevalence rate for this sample was 73% (337/461). It appears that the variability associated with the GAI regarding scores in working memory and processing speed is common in children and adolescents. The 73% prevalence rate was obtained using a <15% base rate. When a <10% base rate is used, the prevalence drops to 55% (255/461). Given these results, caution would need to be given to interpreting working memory and processing speed variability as indicative of clinical importance, and when ascribing clinical importance, it may be prudent to use a more stringent base rate (<10%).

Of those students who met the GAI rules, the two rules met most frequently were rule #2: a significant difference between the PRI and PSI and rule #4: subtest variability within the WMI or PSI. Regarding the PRI-PSI difference, in the majority of cases, the PRI was higher than the PSI. Regarding the variability within the WMI and PSI, Digit Span was less than Letter-Number Sequencing (WMI variability), and Coding was less than Symbol Search (PSI variability). Within the profiles of students who meet the GAI, they are likely to score lower on Digit Span and Coding. It may be that these patterns (PRI>PSI and variability within the WMI and PSI) are more common among non-clinical cases. In addition, although the majority of cases had FSIQ's less than the GAI, consistent with clinical literature, the correlation between the FSIQ and GAI for the 337 students was .93 ($p<.01$) and the mean FSIQ for this subsample was 107 (standard deviation=13), while the GAI mean was 109 (standard deviation=14). This may not be the case in clinical samples.

There are two limitations of this investigation. One involves the sample. The children were selected by graduate students, and although none were diagnosed, there is the potential that some of the children may have disabilities that have not been identified. Thus to improve the non-clinical sample it is recommended that children attending public schools who are performing adequately and have never been referred would be selected. The second limitation is the absence of a clinical sample for comparison purposes. Such a sample could be chosen from a clinical population reflecting several diagnoses (e.g., LD, ADHD) and matched to the non-clinical group based on age, grade, and gender characteristics.

Additional studies may be designed to (a) investigate prevalence among matched controls using other variables (e.g., academic achievement, behavior) to examine what factors correlate with lowered WMI and PSI scores, (b) analyze the patterns of differences among clinical versus non-clinical samples (e.g., VCI-WMI versus PRI-PSI; subtest score differences; mean difference between GAI and FSIQ), and (c) determine if the use of the FSIQ in samples of students who did or did not qualify for special education classification (e.g., learning disabled) was attenuated to the degree that the use of the GAI would have led to a different eligibility decision (e.g., false negative). As the use of the GAI becomes more common in eligibility decision-making, it becomes more critical to understand its properties.

References

Prifitera, A., Weiss, L.G., & Saklofske, D.H. (1998). The WISC-III in context. In A. Prifitera & D.H. Saklofke (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 1-38). New York: Academic Press.

Raiford, S.E., Weiss, L.G., Rolfhus, E., & Coalson, D. (2005). *WISC-IV General Ability Index* (WISC-IV Technical Report No. 4). http://harcourtassessment.com/hai/Images/pdf/wisciv/WISCIVTechReport4.pdf

Sattler, J.M. (2001). *Assessment of Children: Cognitive applications* (4th ed.). San Diego, CA: Author.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition. Administration and scoring manual.* San Antonio, TX: Harcourt Assessment, Inc.