The WISC-IV and Children and Adolescents with Intellectual Disability: Evaluating for Hidden Floor Effects in the US Version

Allyssa Lanza

M.S., Antioch University New England, 2012

B.A., University of New Hampshire, 2008

The research reported herein was performed pursuant to a grant from Policy Research, Inc. as part of the U.S. Social Security Administration's (SSA's) Improving Disability Determination Process Small Grant Program. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of Policy Research, Inc., SSA or any other agency of the Federal Government.

Abstract

This study is a recreation of research by Whitaker and Gordon (2012) that assesses for possible floor effects in the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). Their study suggested that the Index and Full Scale IQ (FSIQ) scores of low IQ adolescents taking the WISC-IV (UK version) were significantly inflated because of low raw scores being converted to scaled scores of 1. Since the WISC-IV is the most commonly used intelligence test in school settings in the United States (Riccio, Houston, & Harrison, 1998), this author assessed for similar findings in the US version of the WISC-IV by completing a pilot replication of their study. Additionally, this author created her own adjusted scoring system that more modestly alters WISC-IV scaled scores. Results of this study suggest that Index and FSIQ are not significantly affected by the alternative scoring systems. The author concludes that this area warrants further research.

Keywords: WISC-IV, intellectual disability, floor effects, intelligence testing, intellectual assessment

The WISC-IV and Children and Adolescents with Intellectual Disability: Evaluating for Hidden Floor Effects in the US version

The Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV), is an intelligence test commonly used to assess for intellectual disorders (Riccio, Houston, & Harrison, 1998). The *Diagnostic and Statistical Manual, Fifth Edition* (DSM-5), encourages a well-rounded psychological assessment, including testing and evaluation of functional impairment, to diagnose intellectual disability (American Psychiatric Association, 2013). Since a intelligence testing can be used to support access disability services (Social Security Administration, 2014, April 3), the WISC-IV indirectly can influence access to extremely helpful resources. In this paper, the author outlines research that raises concerns about the ability of the WISC-IV to assess ID, highlights the potential negative impact of an invalid intelligence test for test-takers with ID, and reviews her study.

Definition of Terms

Intelligence. Wechsler (1939) defined intelligence as the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his or her environment. He elaborated that intelligence is not defined by one single ability, but the composite of multiple abilities. For example, Indexes of the WISC-IV break intelligence into 4 main factors, or Indexes, described in this section under "Index score." Collectively, these 4 factors represent major domains that comprise effective cognitive abilities.

Intellectual Disability (ID). The DSM-5 diagnosis of intellectual disability (also called intellectual developmental disorder) replaced *The Diagnostic and Statistical Manual, Revised Fourth Edition* (DSM-IV-TR; American Psychiatric Association, 200) diagnosis of mental retardation (MR). There are many similarities between the MR and ID diagnoses, but the key difference is that ID heavily focuses on functional impairment rather than intelligence testing.

The first criterion is that a person experiences difficulty in general mental abilities related to “reasoning, problem-solving, planning, abstract thinking, judgment, academic learning and learning from experience confirmed by both clinical assessment and individualized, standardized intelligence testing” (American Psychiatric Association, 2013; pp. 33). To elaborate, intellectual deficits must be both objectively measured by intelligence tests, such as the WISC-IV, but also observed by trained clinical professionals. In other words, a FSIQ less than 70 without observed impairment would not suffice for diagnoses, nor would a person with a FSIQ in the normal range but with challenges in the domains listed above.

The second criterion is that the deficit in mental abilities must significantly affect performance in one or more aspects of daily life, like “communication, social participation, and independent living, across multiple environments, such as home, school, work, and community” (American Psychiatric Association, 2013; pp. 33). This is slightly more descriptive than the criterion for MR, which does not specify any requirement for needed support.

The final required criterion is that onset must occur during the developmental period (American Psychiatric Association, 2013). This is different from the MR criterion that requires that a person have symptoms present as a child or adolescent. The term “developmental period” allows flexibility in the observation of symptoms to young adulthood, where some individuals in the very mild, 65-70 IQ range may display significant difficulties adjusting to independent living. For young adults who may have had their intellectual needs neglected as children and adolescents, this offers them opportunities to be accurately diagnosed retroactively and to be potentially provided support services for people with ID.

Severity of ID is classified as mild, moderate, severe, or profound. Criteria for each ID specifier are based on qualitative information and the individual's functioning (American

Psychiatric Association, 2013). This is different from the DSM-IV's MR severity specifiers, which were based on actual or estimated IQ level (American Psychiatric Association, 2000). Additionally, the DSM-5 has created a separate diagnosis of unspecified intellectual disability for when a person over the age of 5 is unable to be assessed due to physiological or co-morbid disorders that impact assessment. This has replaced the DSM-IV's diagnosis of MR, unspecified.

Raw score. After the administration of each subtest on the WISC-IV, a raw score is calculated based on the scoring criteria in the manual (Wechsler, 2003). Often zero to two points are awarded for each item on a subtest, based on the accuracy of the responses. The sum of the items provides a raw score for each subtest.

Scaled score. The scaling process translates the child's subtest raw score to a standard score that is meaningful when the child is compared to their same-aged peers. This is practical since the raw scores of young children are likely to much lower than older adolescents. The scaled scores range from 0-19 with a mean at 10. Once scaled scores are calculated, a child's performance on each subtest can be easily mapped as at, above, or below the average abilities of other children within that age group. These scores are calculated using a table in the WISC-IV scoring handbook or by scoring software, and the scaled scores were developed based on the test norming sample (Wechsler, 2003).

Index score. Indexes represent a person's relative intellectual strengths and weakness (Flanagan, & Kaufman, 2009). Each Index is comprised of particular subtests, and Index scores are calculated from the standard scores of subtests. The Index scores are used to measure ability, such as Verbal Comprehension Index (VCI), Working Memory Index (WMI), Perceptual Reasoning Index (PRI), and Processing Speed Index (PSI). The VCI is calculated from the subtest scores on Similarities, Vocabulary, and Comprehension (Information and Word Reasoning are supplemental subtests). The PSI is calculated from Block Design, Picture Concepts, and Matrix Reasoning (Picture Completion is a supplementary subtest). The WMI is calculated from Digit Span and Letter-Number Sequencing (with Arithmetic is a supplementary subtest). The PSI is calculated from Coding and Symbol Search (Cancellation is a supplementary subtest). Indexes are thought to have more utility in assessing strengths and weaknesses because subtests are too specific and variable to reliably infer broad abilities.

Full Scale IQ (FSIQ). The FSIQ is a numerical value that represents an individual's general intellectual ability. It is calculated from the Index scores and does not represent any relative strengths or weaknesses. For many with typical intellectual development, this value suffices to generalize intellect because an individual's intellectual strengths and weaknesses tend not to differ greatly (Flanagan, & Kaufman, 2009). In other words, a person with a typically developing intellect may have personal strengths and weaknesses, but generally his or her abilities measured on the Indexes will not significantly differ. The FSIQ can still represent how the person generally performs. For people with ID, FSIQ is less useful because there tends to be greater differences between the abilities represented within an Index score (MacLean et. al., 2011). Additionally, differences between subtests are likely to be significant. By using scores that generalize multiple abilities, the significant differences in strengths and weaknesses are lost in an averaged value.

Floor effects. This term represents a phenomenon that occurs when a test is unable to measure below a particular value. A common result of floor effects is that an examinee obtains an inaccurate, higher score. Examples of how floor effects could occur include not having enough “easy” items on a subtest so that the examinee can meet an appropriate baseline, or when there are not enough “easy” items to describe the examinee's abilities to perform on the subtest. Hidden floor effects refer to floor effects that are not necessarily obvious to an examiner. An example of a hidden floor effect would be if the items required to meet baseline were

significantly easier than the later items on a subtest. Although the examinee may be able to meet baseline criteria, the items of the test still measure beyond the abilities of the examinee.

Wechsler Intelligence Tests

Standardized tests are necessary to diagnose intellectual disorders in the DSM-5 (2013) except under extenuating circumstances. Intelligence tests are frequently used to inform treatment because they may identify strengths and a weakness in a student's learning style. Thus, the assessment measures used to assess for ID must be studied and critiqued for their validity with the ID population.

Wechsler Four Factor Model

The Wechsler intelligence tests were the primary measures for assessing intellectual ability for many years. Published in 1939 by David Wechsler, the Wechsler-Bellevue Intelligence Scale was the first of the Wechsler IQ tests (Wechsler, 2003). At the time, they were developed without theory, and Wechsler believed that tests gave insight into a client's personality. Since then, Wechsler-based assessment tests have been revised to incorporate a four factor model of understanding intelligence and are used internationally. The most common two Wechsler tests are described here.

The Wechsler model theorizes that the Indexes represent the four main domains of intelligence (Wechsler, 2003). The Verbal Comprehension Index (VCI) measures verbal knowledge and comprehension, and is often seen as a good predictor of scholastic achievement. The Perceptual Reasoning Index (PRI) measures fluid reasoning and perceptual and organizational skills. Fluid Reasoning is the ability to apply learned skills to novel or unfamiliar situations, which often utilizes perceptual and organizational skills. The Working Memory Index (WMI) measures short-term auditory memory, concentration, and attention. Finally, the Processing Speed Index (PSI) measures the speed that one processes nonverbal visual information.

Standardization of the WISC-IV

The WISC-IV is an intelligence test for children and adolescents aged 6 years to 16 years, 11 months (Wechsler, 2003). It purports to measure intellect from $40 \leq FSIQ \leq 160$. Norms for the test were developed in a five-stage process. The Standardization Phase used a stratified sample of 2,200 children, with 200 samples per age bracket. Samples were identified using trained recruiters and independent examiners. Some children were excluded from the study, including those who were taking medication that might depress performance, such as antipsychotics and antidepressants. Approximately 5.7% of the norming sample was added to "accurately represent the population of children attending school" (Wechsler, 2003, p. 23). No further information was provided to clarify the demographics of the 5.7% or children attending school.

Despite this information, the Wechsler Technical and Interpretive Manual (*Manual*, Wechsler, 2003) posited that the sample was likely affected by selection bias. It reported children in the sample were relatively high-functioning and that hospitalized children were excluded from the study. Additionally, this author argues that excluding children on medication likely affected the sample due to the high prevalence of children with MR/ID on medication. Depending on the abilities and functioning of children in the WISC-IV sample, it is possible that their sample was not highly representative of children and adolescents with MR/ID.

The author also posits that Wechsler is minimizing the variability of Index scores within each sample. Although the standard deviations for each Index may be less than the general population, Wechsler (2003) found that 16.7% of children with FSIQ less than 79 points had PRI scores 15 or more points higher than VCI scores, and 10.2% of children in this range had VCI scores 15 or more points higher than their PRI scores. Thus, at least 26.9% of children with FSIQ

scores below 79 have Index scores that differ by 15 or more points. This suggests to the author that variance is likely not occurring within each Index, but between the Indexes. No information on the prevalence of invalid FSIQ scores was reported in the *Manual*.

Evaluating Floor Effects in the UK Version of the WISC-IV

Whitaker and Gordon (2012) researched Whitaker's (Whitaker, 2008; Whitaker, 2010; Whitaker, & Wood, 2008) hypothesis that scaled scores of 1 create a hidden floor effect. To measure their hypothesis, they created adjusted scaled scores for raw scores that otherwise would have been scaled to 1. From the scoring tables found in the *Manual* (2003), they extrapolated the algorithm the Wechsler tests used to distribute scaled scores less than or equal to 10. They then applied the algorithm to very low raw scores, creating adjusted scores of 0 and below. They calculated Indexes and FSIQ based on the using the adjusted scores to see if, and to what extent, they varied.

Whitaker and Wood found that 45 out of 66 raw scores that would have been scaled to 1 qualified for a lower scaled score using their method. Furthermore, nine out of 17 subjects had a reduced FSIQ after their scores had been adjusted. Of these nine scores, four were within six points of the original FSIQ, and five had a greater reduction. These change in these scores created a significantly different score distribution, and Whitaker and Wood posited that the change in score distribution suggested a floor effect in the UK version of the WISC-IV.

Resources for People with ID

Recent changes in legislation have provided people with ID numerous legal rights. The Americans with Disabilities Act (ADA; 1990) and the ADA Amendments Act of 2008 (2008), provided rights to people with ID, such as equal pay and access to public entities. The Acts were passed to promote equality and diminish discrimination against individuals with physical and mental disabilities. Subsequently, the Acts have allowed individuals with ID access to specialized education and social services, such as financial disability benefits.

To be covered by the Acts, a person must fall under one of these criteria: "(A) physical or mental impairment that substantially limits one or more major life activities of such individual; (B) a record of such an impairment; or (C) being regarded as having such an impairment" (ADA, 2008). When compared to the diagnostic criteria for both MR and ID, as defined above, it appears that all people who qualify for MR (and likely ID) will also be eligible for the ADA.

Supplemental security income (SSI) through the Social Security Administration (SSA) is another valuable resource for individuals with ID. In order to qualify for SSI, one must meet criteria and program requirements. There are a number of ways a person may meet disability criteria for SSI, which are highlighted in their publications online. The medical criteria for Intellectual Disability eligibility are detailed in section 112.05 (Social Security Administration, 2014, April 2; available in Appendix A). Per their definition, a disability is defined as the "inability to engage in any substantial gainful activity (SGA) by reason of any medically determinable physical or mental impairment(s)...." (Social Security Administration, 2014, April 4). These impairments must impact the person or child for longer than one year, and meet requirements determined in a sequential evaluation process. The sequential evaluation process for children and adolescents (aged 22 or less) includes a "review of the child's current work activity (if any), the severity of his or her impairments(s), and an assessment of whether his or her impairment(s) results in marked and severe functional limitations" (Social Security Administration, 2014, April 3).

Statement of the Problem

Clinical presentations of people with ID vary significantly, and their needs for educational, occupational, and social success are highly dependent on their unique strengths and weaknesses. Intellectual testing is commonly used to identify overall cognitive capacity, as well

as the particular ways an individual learns and processes information. Information gathered from intelligence testing, such as that provided from the WISC-IV, informs what services are appropriate for children and adolescents with ID. Inflation of abilities leads to a disservice for children and adolescents requiring support and assistance, and can have lasting consequences in their ability to access services as adults.

For over a century, intelligence testing has been used to measure people's ability to process information (Wechsler, 2003). Many intelligence tests have been revised to suit changing definitions of intelligence and re-normed to represent changes in the population's cognitive abilities. Currently, intelligence tests are commonly used not only to assess overall intellectual ability, but to inform individualized education programs and aid in differential diagnosis.

The rationale for the study was based on problems identified in Wechsler tests and the significance of an ID diagnosis. There is ample research that older and UK editions of Wechsler intelligence tests have not accurately assessed low intellect (MacLean, McKenzie, Kidd, Murray, & Schwannauer, 2011; Whitaker, 2008; Whitaker, 2010; Whitaker, & Wood, 2008); However, there is a lack of research about the US version of the WISC-IV. There is insufficient evidence from non-Wechsler-funded research that supports the WISC-IV's abilities to assess low IQ. Due to the significance of an ID diagnosis, and the prevalence of Wechsler tests in determining ID, it is imperative that research explore these concerns.

Research suggests that the specific abilities of people with ID vary, meaning that there are larger differences between strengths and weaknesses in people with ID than in people with average intellect. MacLean, McKenzie, Kidd, Murray, and Schwannauer (2011) found that the WAIS-III Index scores over-generalize the large range of abilities represented by subtest scores. While the specific abilities of people with typically developing intellect often cluster around an Index score value, one Index score may not best represent the varied abilities of a person with ID. The authors argued that other Wechsler intelligence tests may also over-generalize because they also use Indexes.

Wechsler tests have been criticized for insufficient norm samples for low IQ (Whitaker, 2008; Whitaker, 2010; Whitaker, & Wood, 2008). Since norming samples are used to standardize the test and to create scores that meaningfully compare an individual's results to the general population, such an oversight posed many problems. In particular, an insufficient norming sample is unlikely to represent the abilities of the population in that range. A small sample selection increases the likelihood that anomalies of the group will be generalized to the entire population. A larger sample decreases the likelihood that the sample group will be inappropriately homogeneous, that less common traits will be assumed to be more prevalent in the population, or both. Although the norming sample for children with low IQ has been expanded, Wechsler notes problems with sample selection bias.

Wechsler (2003) reported that the increased norming sample for children with low IQ improved the WISC-IV's ability to assess variability within Indexes; however, further verification is necessary. This author was unable to find research that corroborates that variability between Indexes has improved. Due to the large differences between the Indexes, FSIQs of children with low cognitive abilities may overgeneralize their highly varied strengths and weaknesses.

Additionally, undetected floor effects have been hypothesized in Wechsler tests. Floor effects occur when a score, such as a FSIQ, cannot accurately measure below a particular value. Wechsler (2003) states that the WISC-IV's floor is at a FSIQ of 40. Studies of UK versions of the WAIS-III, WAIS-IV, and WISC-III suggested floor effects may occur as high as FSIQ of 70 (Whitaker, 2008; Whitaker, 2010; Whitaker & Wood, 2008).

Whitaker and Wood (2008) posited two reasons that there are floor effects. The first

reason this may occur is because low raw scores are scaled to scores of 1, including raw scores of zero. The second reason is because the distribution of intellect in the population is assumed to be normal bell-shaped. Whitaker (Whitaker, 2008; Whitaker, 2010; Whitaker & Wood, 2008) argued that the scaled score of 1 represents both children who perform extremely poorly and children who cannot perform the task at all. As a result, he hypothesized that the scaled score of 1 becomes meaningless because it does not differentiate between severe low ability and no ability to complete subtest tasks.

Whitaker and Wood's research (2008) suggested that due to the high number of scaled scores of 1 received by people with FSIQ below 70, it is possible FSIQ scores are inflated. For children and adolescents with FSIQ scores in the 60s and 70s, a few lower points may be the deciding factor for a disability diagnosis or disability services. They hypothesized that more test-takers would get FSIQ scores below 70 if the Wechsler scoring system (2003) used a scoring method that did not scale extremely low raw scores to scaled scores of 1.

In addition, Whitaker (Whitaker, 2008; Whitaker, 2010; Whitaker & Wood, 2008) expressed concern that Wechsler tests assumed a normal distribution in a population's intellect. Instead, Whitaker argued that intellect is likely bimodal. He cited the increase in disorders affiliated with low intellect (such as autism spectrum disorders). He suggested that an assumed normal curve may affect the standardization of the test. The *Manual* (Wechsler, 2003) does not include the distribution of FSIQ scores collected from their standardization sample. The *Manual* refers to collecting stratified samples based on age, sex, race, parent education level, and geographic location (pp. 20-21). Due to the limited psychometric information in the *Manual*, it is unclear if and how an assumed distribution affects low IQ scores on the WISC-IV.

Finally, the study was important due to the significance of an ID diagnosis and the role intelligence testing plays in ID assessment. The DSM-5 diagnostic criteria for ID include: deficits in general cognitive abilities, significant problems in functioning as a result of cognitive deficits, and onset during the developmental period (American Psychiatric Association, 2013). The DSM-5 encourages diagnosis not to be based entirely on intelligence testing, but instead on a thorough psychological evaluation in conjunction with testing. Consistent with MR diagnosis in the DSM-IV, ID would be suggested by an intelligence quotient (IQ) score at least two standard deviations below the mean (at or below a FSIQ of 70 on the WISC-IV; Wechsler, 2003).

The present author argued that while diagnosis of intellectual disability will be improved when testing is used in conjunction with assessment; she has concerns that elevated FSIQ and Index scores may mislead an assessor to miss ID diagnoses. She argued that the WISC-IV must accurately measure low intellect not only to accurately inform assessors since it is trusted as a reliable measure of intelligence. She posits that if IQ tests inflate cognitive abilities, that impairments in functioning may be understood as oppositionality, unwillingness, or general laziness.

The range of FSIQ below 70 was selected because individuals with these scores are at increased risk for being misdiagnosed. The author hypothesized that children and adolescents with very low intellect were at lesser risk for being misdiagnosed due to the severity of their limitations and because of revisions to the DSM-5 (2013) that emphasized functional impairment. The author argued that slight inflation in FSIQ scores below 70 may lead some assessors and support service agencies to deny diagnosis of ID since functional impairment would be less overt than in lesser IQs. As a result, people in this range are for increased risk to miss criteria for ID and subsequent services.

Research Questions

The study answered the following questions:

- Are hidden floor effects hypothesized in the UK version of the WISC-IV present in the

US version?

- Is there evidence that scaled scores of 1 significantly inflate Index and FSIQ scores?
- Is there a significant difference in the number of children and adolescents who might qualify for a diagnosis of ID when a scaled score of zero is utilized?
- Does Lanza's adjusted scores produce different results than Whitaker and Gordon's (2012)?

Research Hypotheses

The research hypotheses for the study were as follows:

1. No hidden floor effects will be found in this sample.
2. There will be no significant difference in Index or FSIQs.
3. There will be no difference in results based on adjusted scoring method.

Method

The author used qualitative research methods to measure for floor effects in the WISC-IV when assessing children and adolescents with FSIQs below 70. The author used the adjusted scoring system proposed by Whitaker and Gordon (2012) to assess if the U.S. WISC-IV scoring system inflates the Index and FSIQ scores of children with low IQs.Whitaker and Gordon (2012) began creating in England their adjusted scoring system by using the data available in the scoring charts of *Wechsler Intelligence Scale for Children–Fourth Edition: Administration and Scoring Manual* (2003). Since the WISC-IV does not provide the equation they used to determine how raw scores would be converted to scaled scores, Whitaker and Gordon (2012) found an algorithm by plotting the mathematical relationship between the raw scores and scaled scores less than 10. Only scaled scores less than 10 were included because they felt the mathematical relationship would be simpler when only low scores were used. They observed that line did not continue as expected with very low raw scores, and that instead it stopped abruptly at the scaled score of 1. They hypothesized that this represented the suspected floor effect. Figure 1 (found in Appendix B) is a visual representation of one of the graphs created using the Wechsler raw to scaled score tables.

Whitaker and Gordon (2012) created adjusted scaled scores by allowing the relationship between raw scores and scaled scores to continue below a scaled score of 1. For example, there is a linear mathematical relationship between raw scores and scaled scores on the subtest Digit Span for children aged 7 years, 8 months to 7 years, 11 months. Figure 1 in Appendix B shows that a straight line and a linear equation best fit the points provided by the WISC-IV conversion charts. Table 1 in Appendix C clearly shows a 1:1 linear relationship between raw and scaled scores until it reaches low raw scores.

Whitaker and Wood (2012) continued the mathematical relationship found via the scores in the WISC-IV *Manual* (2003) to the very low raw scores. Where the WISC-IV does not allow scaled scores to go below 1, their adjusted scores did not have a lower limit. Table 2 in Appendix D shows how the scores were adjusted for Digit Span, ages 7 years, 8 months to 7 years, 11 months.

Whitaker and Wood (2012) found that Index and FSIQ scores were significantly lower when using the adjusted scores. Since the UK version of the WISC-IV is different than the U.S. version, the author investigated whether similar results could be seen in the U.S. version of the WISC-IV. To do so, she replicated the aforementioned study and created another method of adjusting scores.

Setting

Historical data were collected from assorted schools located in New Hampshire and Massachusetts. One set of data was collected from a southwestern New Hampshire school district that serves students from over 200 square miles, 14 school buildings, and approximately

4,200 students. The other set of data was collected from a suburban Massachusetts school for students with emotional, behavioral, and developmental disorders.

Participants

The sample consisted of seven students who scored a FSIQ below 70 on the WISC-IV. Students were identified by their schools as needing intelligence testing for numerous reasons including, but not limited to, admission, overall poor academic achievement, concerns about particular areas of learning, or unexplained conduct problems. Participants ranged from the ages of 7 years, 8 months to 15 years, 10 months at the time of testing. Gender, race, and socioeconomic status were not known. Table 3 in Appendix E lists the ages and locations of each participant.

Measures

Students were administered the Wechsler's Intelligence Scale for Children, fourth edition (WISC-IV; Wechsler, 2003). Participants had been administered the 10 core WISC-IV subtests as specified in the administration manual (Wechsler, 2003).

Procedures

Per APA's ethical codes (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), each school gave permission to access archived data of children and adolescents without parental consent because the data were owned by the schools for administrative reasons. By giving permission to access their test data, the schools received the findings of the study.

In order to protect the identities of the students whose tests were used in the study, all WISC-IV protocols were de-identified prior to coming into the author's possession. Physical protocols were linked with electronic data using randomly assigned numbers. Raw, scaled, Index, and FSIQ scores were recorded from the WISC-IV protocols into a password-protected Microsoft Excel file. The hard copies of de-identified assessment protocols were stored in a locked filing cabinet at the author's residence.

Data Analyses

The methods of analysis were informed by the research of Whitaker and Gordon (2012), and done in support of Wright State University's Statistical Consulting Center. Two sets of adjusted scores were created for subtests receiving a scaled score of 1. The first set was made using the exact method of Whitaker and Gordon (2012). The second set (referred heretofore as Lanza's adjusted scores) was informed by Whitaker and Gordon (2012), but was modified slightly.

Whitaker and Gordon's adjusted scores were created by finding the mathematical relationship between the WISC-IV (U.S. version) raw scores to scaled scores for each subtest receiving a scaled score of 1. It is important to note that many subtests have a range of raw scores that convert to the same scaled score. In this instance, Whitaker and Gordon took the highest score of the raw score range to graph (personal communication). For example, if raw scores 0 to 3 received a scaled score of 1, Whitaker and Gordon (2012) plotted the relationship between a raw score of 3 and a scaled score of 1.

This author modified their method slightly and used the mean of the raw scores in the range when calculating the mathematical relationship between raw scores and scaled scores. The author considered that the highest raw score might not best represent the scaled score. Instead, she used the average raw score to correlate with each scaled score in order to represent the mean raw score value of each scaled score. For example, if raw scores 0 to 3 received a scaled score of 1, the author plotted the relationship between a raw score of 2 (the integer closest to the mean of raw scores 0, 1, 2, and 3) and a scaled score of 1.

The raw and scaled scores were graphed using Microsoft Excel for both Whitaker and Gordon and Lanza's adjusted scores. Using regression analysis, Microsoft Excel identified the slope of the line that best fit the plotted points. This provided an algorithm for where the line would continue should it be allowed to continue past a scaled score of 1. Index and FSIQ scores were then calculated using the new scaled scores according to Wechsler scoring system and norms tables. Table 4 in Appendix F lists the formulae found for each subtest requiring an adjusted score.

Adjusted scaled scores were created for subtests on which a student obtained a scaled score of 0. Adjusted scaled scores for other subtests were not calculated since they would not be used in the study sample. Although the Lanza adjusted score equations were all different from the Whitaker and Gordon adjusted score equations, it is important to note that this did not necessarily create differences in the two adjusted scoring systems. This is discussed further in the results section.

Adjusted scores were then added together to create the Sum of Scaled Scores for each Index and for the FSIQ. Occasionally, new Index scores could not be calculated using the Wechsler scoring charts because they were too low. When this happened, new Index scores were created using a very similar process as the creation of adjusted scores. The mathematical relationship between the sums of the scaled scores was plotted with the Index scores to produce an equation. The equation was then used to predict what the Index score would be if lower scaled scores had been available in the original Wechsler scoring.

Once new FSIQ and Index scores were created for each sample, the Wechsler scores and adjusted scores were compared using a two-tailed paired-sample t-test. In order to determine differences between the Wechsler and both the Lanza and Whitaker & Gordon (2012) adjusted scores, the data were analyzed using Microsoft 10 Excel and Statistical Package for the Social Sciences, version 20 (SPSS 20). Both Whitaker and Gordon's and Lanza's adjusted scores were run as "post" scores to see if both score adjustment procedures created significant difference in the FSIQ and Index scores.

Results

Analyses for this study were performed on seven WISC-IV protocols of students. Adjusted scores were created using the same process as Whitaker and Gordon (2012). Additional consultation was sought from direct communication with Simon Whitaker to better understand the methodology of Whitaker and Gordon (2012). Analyses for the study were done with the support of Wright State University's Statistical Consulting Center.

Rescoring Protocols Using Adjusted Scores

Each protocol was rescored using Whitaker and Gordon Adjusted Scores and Lanza Adjusted Scores. Out of the seven total protocols used in this study, a total of four had changes to their Index scores using an adjusted scoring system. Two had had changes in Index scores using Lanza's adjusted scores, and four had changes using Whitaker and Gordon's (2012) adjusted scores. One of the protocols had no difference in scores with either method. A total of three protocols had changes to their FSIQ using an alternative scoring method. The Wechsler, Whitaker and Gordon adjusted, and Lanza adjusted scores are listed in the tables 5 through 11 in Appendix G.

Testing Differences in Index and FSIQ Scores

Ten paired *t*-tests were used to determine if the Whitaker and Gordon (2012) or Lanza adjusted Index and FSIQ scores were significant differently from the Wechsler Index and FSIQ scores. There were no significant differences noted between Indexes or FSIQ between the Wechsler scores or the Whitaker and Gordon (2012) Adjusted or Lanza Adjusted. Table 12 in Appendix H shows the results of the *t*-tests.

Based on the paired-samples *t*-tests, neither Whitaker and Gordon's (2012) method of rescaling nor this author's method of scoring created significant changes to the sample's FSIQ. Furthermore, it should be noted that the adjusted FSIQ scores that were affected by the alternative scoring system still fell within the Wechsler predicted FSIQ range with 95% confidence. Only the protocol of Child 3 had Whitaker and Gordon (2012) and Lanza adjusted FSIQ scores outside of the FSIQ range expected with 90% confidence. Table 13 (found in Appendix I) shows the FSIQ ranges for each protocol to highlight how the adjusted scores do, or do not, fall within the expected ranges.

Summary and Conclusions

This study evaluated for a hidden floor effect in the US version of the WISC-IV. Evidence supports that the Wechsler four factor model poses problems to assessing the IQ of children and adolescents with ID, although no study exists specifically that evaluates the US version of the WISC-IV. Findings from this study were limited due to the small sample; however, the research cites and method used in this study can inform SSA on how it uses intelligence testing and how it may better understand the limits of assessment measures.

Both Whitaker and Gordon's (2012) and Lanza's adjusted scores affected the subtest scores of the protocols used. Five paired-sample *t*-tests were conducted to determine differences between the Wechsler scores and Lanza adjusted scores, and five pair-sample *t*-tests were conducted to determine differences between Wechsler scores and Whitaker and Gordon (2012) scores. There was no significant difference found in the Index scores or FSIQs of either adjusted scoring system.

Results of this study must be interpreted cautiously for many reasons. The first reason is that this study had a limited data. Not only was it less than half of Whitaker and Gordon's (2012) study, but data sources were limited to two school locations. Protocols for this study were provided by individuals from two New England schools. The sample of students represents an extremely narrow part of the population. Based on demographics of the surrounding areas of the schools, students were likely to be Caucasian and from middle to lower socioeconomic status. If this study were to be expanded or replicated, it would be useful to gather a larger number of protocols from various sources, socioeconomic status, race, and regions within the United States. All results of the study should be noted to be found only in students with similar demographics. Results should not be assumed to be present in other populations. The small sample may inflate or miss any findings that may be found in a larger sample. Secondly, any claims that the WISC-IV is not a valid measure of cognitive abilities for low intellect has many implications on its use in diagnosis of intellectual disability and allocation of funds.

This study is heavily based on theory that is challenging to research. Many of Wechsler's arguments were theoretical, and testing his theories would require broad-scale data collection. For example, an international study would need to be conducted to see if his hypothesis that intellect is bimodal was supported. Another example would be collecting a new norming sample for the WISC-IV to test the hypothesis that the existing sample is inadequate. Since these studies do not exist at this time, smaller studies such as this can only allude to support of the larger hypotheses.

Finally, the sample collected for this study did not control for variables that have been controlled for in the WISC-IV norming sample. Specifically, this study did not determine if the child or adolescent was on pharmacological medication at the time of testing. This author argues that while some medication might influence WISC-IV performance, such as creating sedating effects or increasing a child's ability to focus, children in this sample were more representative of the population of children and adolescents with ID.

This area of research would benefit from an expanded version of this study to see if, and

what extent, other floor effects are observed using both Whitaker and Gordon's (2012) and this author's adjusted scoring system. Due to the significant results observed in the study by Whitaker and Gordon (2012), it is important to understand if their findings are unique to the UK version of the test or because of their larger sample.

In conclusion, there was no evidence suggesting a possible floor effect on the Processing Speed Index of the US Version of the WISC-IV. This author hopes that assessors will be mindful in interpreting scaled scores of 1 on the WISC-IV, and that they will critical of very low raw scores being scaled to 1. She hopes that assessors might regularly use corroborating measures when there are numerous scaled scores of 1 on a WISC-IV protocol, and that they might consider including a discussion paragraph in their reports discussing the possibility of floor effects on the test. Finally, she hopes that assessors embrace the spirit of the DSM-5 Intellectual Disability diagnosis, and focus their assessments on describing how an individual is likely to function given their intellectual abilities. She believes that this will allow children and adolescents with ID to receive access to services they require.

References

ADA Amendments Act of 2008, Pub. L. 110–325, 122 Stat. §3553 (2008).

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing.* Washington, D.C.: American Educational Research Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (Revised 4th ed.)*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5 (5th ed.)*. Washington, DC: Author.

Americans with Disabilities Act of 1990, Pub. L. No. 101-336 (1990).

Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC-IV assessment (2nd ed.).* Hoboken, NJ: John Wiley & Sons.

Individuals with Disabilities Education Act (IDEA) Amendments of 1997, Pub. L. No. 105-17.

Individuals with Disabilities Education Improvement Act (IDEIA) Pub. L. 108-446 (2004).

MacLean, H., McKenzie, K., Kidd, G., Murray, A. L., & Schwannauer, M. (2011). Measurement invariance in the assessment of people with an intellectual disability. *Research in Developmental Disabilities*, *32*(3), 1081-1085. doi:10.1016/j.ridd.2011.01.022

McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll Cognitive-Achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools, 47*(7), 651-675.

Newton, J. H., & McGrew, K. S. (2010). Introduction to the special issue: Current research in Cattell–Horn–Carroll–based assessment. *Psychology in the Schools, 47*(7), 621-634.

Riccio, C. A., Houston, F., & Harrison, P. L. (1998). Assessment practices for children with severe mental retardation. *Journal Of Psychoeducational Assessment*, *16*(4), 292-301. doi:10.1177/073428299801600401

Social Security Administration. (2014, April 3). http://www.ssa.gov/disability/professionals/bluebook

Social Security Administration. (2014, April 2) http://www.ssa.gov/disability/professionals/bluebook/112.00-MentalDisorders-Childhood.htm#112_05

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD US: Williams & Wilkins Co. doi:10.1037/10020-000

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children – Fourth Edition: Administraion and Scoring Manual.* San Antonio, TX: Pearson.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition*. London: Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children – Fourth Edition: Administration and Scoring Manual.* San Antonio, TX: Pearson.

Wechsler, D. (2008a). *Wechsler Adult Intelligence Scale–Fourth Edition*. San Antonio, TX: Pearson.

Wechsler, D. (2008b). *Wechsler Adult Intelligence Scale–Fourth Edition: Technical and interpretive manual*. San Antonio, TX: Pearson.

Whitaker, S. (2008). WISC-IV and low IQ: Review and comparison with the WAIS-III. *Educational Psychology in Practice, 24*(2), 129-137. doi:10.1080/02667360802019180

Whitaker, S. (2010). Error in the estimation of intellectual ability in the low range using the WISC-IV and WAIS-III. *Personality & Individual Differences, 48*(5), 517-521. doi:10.1016/j.paid.2009.11.017

Whitaker, S., & Wood, C. (2008). The distribution of scaled scores and possible floor effects

on the WISC-III and WAIS-III. *Journal of Applied Research in Intellectual Disabilities, 21*(2), 136-141. doi:10.1111/j.1468-3148.2007.00378.x

Whitaker, S., & Gordon, S. (2012). Floor effects on the WISC-IV. *International Journal of Developmental Disabilities, 58*(1), 1-9. doi: 10.1179/2047387711Y.0000000012

Yalon-Chamovitz, S. (2009). Invisible access needs of people with intellectual disabilities: A conceptual model of practice. *Intellectual & Developmental Disabilities, 47*(5), 395-400

Appendix A: SSA Requirements for 112. 05 Intellectual Disability

Characterized by significantly subaverage general intellectual functioning with deficits in adaptive functioning.

The required level of severity for this disorder is met when the requirements in A, B, C, D, E, or F are satisfied.

A. For older infants and toddlers (age 1 to attainment of age 3), resulting in at least one of the appropriate age-group criteria in paragraph B1 of 112.02; or, for children (age 3 to attainment of age 18), resulting in at least two of the appropriate age-group criteria in paragraph B2 of 112.02;

or

B. Mental incapacity evidenced by dependence upon others for personal needs (grossly in excess of age-appropriate dependence) and inability to follow directions such that the use of standardized measures of intellectual functioning is precluded;

or

C. A valid verbal, performance, or full scale IQ of 59 or less;

or

D. A valid verbal, performance, or full scale IQ of 60 through 70 and a physical or other mental impairment imposing an additional and significant limitation of function;

or

E. A valid verbal, performance, or full scale IQ of 60 through 70 and:

1. For older infants and toddlers (age 1 to attainment of age 3), resulting in attainment of development or function generally acquired by children no more than two-thirds of the child's chronological age in either of paragraphs B1a or B1c of 112.02; or

2. For children (age 3 to attainment of age 18), resulting in at least one of paragraphs B2b or B2c or B2d of 112.02;

or

F. Select the appropriate age group:

1. For older infants and toddlers (age 1 to attainment of age 3), resulting in attainment of development or function generally acquired by children no more than two-thirds of the child's chronological age in paragraph B1b of 112.02, and a physical or other mental impairment imposing an additional and significant limitation of function;

or

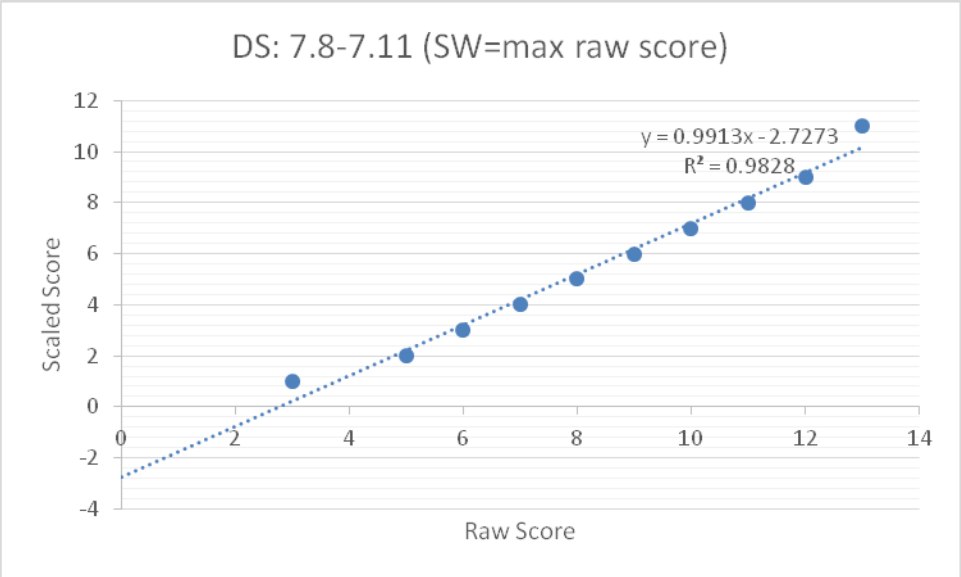
2. For children (age 3 to attainment of age 18), resulting in the satisfaction of 112.02B2a, and a physical or other mental impairment imposing an additional and significant limitation of function.

* Note: Guidelines taken directly from SSA website (Social Security Administration, 2014, April 2)

Appendix B: Figure 1

Figure 1

Relationship Between Raw and Scaled Score on Digit Span for Ages 7:8 to 7:11



Appendix C: Table 1

Table 1
Raw to Scaled Score Conversion Guide for Digit Span, ages 7:8 to 7:11

Raw Scores	Scaled Scores
0	1
1	1
2	1
3	1
4	2
5	2
6	3
7	4
8	5
9	6

Appendix D: Table 2

Table 2

Whitaker and Wood's (2013) Raw to Scaled Score Conversion Guide for Digit Span, ages 7:8 to 7:11

Raw Scores	Scaled Scores
0	-1
1	0
2	0
3	1
4	2
5	2
6	3
7	4
8	5
9	6

Appendix E: Table 3

Table 3
Ages and Location of Each Child

ID	Age	Location
Child 1	15 years, 10 months	New Hampshire
Child 2	11 years, 8 months	New Hampshire
Child 3	7 years, 8 months	Massachusetts
Child 4	9 years, 4 months	Massachusetts
Child 5	12 years, 7 months	Massachusetts
Child 6	8 years, 8 months	Massachusetts
Child 7	8 years, 9 months	Massachusetts

Appendix F: Table 4

Table 4
Formulae for Creating Adjusted Scaled Scores

Subtest	Age Group	Algorithm for Whitaker and Gordon's Adjusted Scores	Algorithm for Lanza's Adjusted Scores
Vocabulary	11:8 to 11:11	$y = 0.3636x - 3.5542$ $R^2 = 0.9983$	$y = 0.3701x - 3.4377$ $R^2 = 0.997$
Comprehension	9:4 to 9:7	$y = 0.5687x - 1.7222$ $R^2 = 0.996$	$y = 0.5824x - 1.6928$ $R^2 = 0.9972$
Comprehension	11:8 to 11:11	$y = 0.5539x - 3.5285$ $R^2 = 0.997$	$y = 0.5558x - 3.3653$ $R^2 = 0.9988$
Matrix Reasoning	11:8 to 11:11	$y = 0.0021x^2 + 0.4762x - 1.969$ $R^2 = 0.9981$	$y = 0.0015x^2 + 0.5012x - 1.9635$ $R^2 = 0.9988$
Digit Span	7:8 to 7:11	$y = 0.9913x - 2.7273$ $R^2 = 0.9828$	$y = 0.9727x - 2.5217$ $R^2 = 0.9832$
Digit Span	11:8 to 11:11	$y = x - 6$ $R^2 = 1$	$y = x - 6$ $R^2 = 1$
Letter-Number Sequencing	11:8 to 11:11	$y = 0.0321x^2 + 0.0688x - 0.346$ $R^2 = 0.9967$	$y = 0.027x^2 + 0.1737x - 0.7363$ $R^2 = 0.9966$
Coding	8:8 to 8:11	$y = 0.2941x - 1.4103$ $R^2 = 0.9962$	$y = 0.2945x - 1.0813$ $R^2 = 0.9985$
Coding	11:8 to 11:11	$y = 0.2687x - 4.3091$ $R^2 = 0.9984$	$y = 0.0031x^2 + 0.0345x + 0.1833$ $R^2 = 0.9948$
Coding	15:8 to 15:10	$y = 0.2248x - 6.0339$ $R^2 = 0.9988$	$y = 0.2315x - 6.0166$ $R^2 = 0.9968$
Symbol Search	11:8 to 11:11	$y = -0.004x^2 + 0.6109x - 3.342$ $R^2 = 0.9966$	$y = -0.0068x^2 + 0.7129x - 3.96$ $R^2 = 0.9966$
Symbol Search	12:4 to 12:7	$y = -0.0029x^2 + 0.577x - 3.9768$ $R^2 = 0.999$	$y = -0.0044x^2 + 0.6383x - 4.3093$ $R^2 = 0.9995$
Symbol Search	15:8 to 15:10	$y = -0.0044x^2 + 0.6334x - 6.9781$ $R^2 = 0.9986$	$y = -0.0059x^2 + 0.716x - 7.7435$ $R^2 = 0.9993$

Appendix G: Tables 5-11

Table 5

Child 1 Wechsler, Whitaker and Gordon Adjusted, and Lanza Adjusted Subtest Scores

Subtest/Index	Raw Scores	Wechsler Scores	Whitaker & Gordon Adjusted Scores	Lanza Adjusted Scores
VCI		81	81	81
Vocabulary	28	4	4	4
Similarities	24	8	8	8
Comprehension	27	8	8	8
PRI		75	75	75
Block Design	30	6	6	6
Picture Concepts	18	8	8	8
Matrix Reasoning	15	4	4	4
WMI		68	68	68
Digit Span	11	3	3	3
Letter-Number Sequencing	15	6	6	6
PSI		50	45*	45*
Coding	23	1	-1*	-1*
Symbol Search	14	1	-1*	-1*
FSIQ		63	60*	60*

Note. * Highlights a difference in score

Table 6

Child 2 Wechsler, Whitaker and Gordon Adjusted, and Lanza Adjusted Subtest Scores

Subtest/Index	Raw Scores	Wechsler Scores	Whitaker & Gordon Adjusted Scores	Lanza Adjusted Scores
VCI		53	47 *	53
Vocabulary	11	1	0*	1
Similarities	7	4	4	4
Comprehension	7	1	0*	1
PRI		57	57	57
Block Design	18	5	5	5
Picture Concepts	10	3	3	3
Matrix Reasoning	6	1	1	1
WMI		50	42*	40*
Digit Span	4	1	-2*	-2*
Letter-Number Sequencing	0	1	0*	-1*
PSI		50	42*	40*
Coding	16	1	0*	2*
Symbol Search	6	1	0*	0*
FSIQ		43	40*	40*

Note. * Highlights a difference in score

Table 7

Child 3 Wechsler, Whitaker and Gordon Adjusted, and Lanza Adjusted Subtest Scores

Subtest/Index	Raw Scores	Wechsler Scores	Whitaker & Gordon Adjusted Scores	Lanza Adjusted Scores
VCI		75	75	75
Vocabulary	13	5	5	5
Similarities	10	8	8	8
Comprehension	7	4	4	4
PRI		67	67	67
Block Design	3	3	3	3
Picture Concepts	8	6	6	6
Matrix Reasoning	7	5	5	5
WMI		65	65	65
Digit Span	3	1	1	1
Letter-Number Sequencing	8	7	7	7
PSI		75	75	75
Coding	30	5	5	5
Symbol Search	18	6	6	6
FSIQ		64	64	64

Note. * Highlights a difference in score

Table 8

Child 4 Wechsler, Whitaker and Gordon Adjusted, and Lanza Adjusted Subtest Scores

Subtest/Index	Raw Scores	Wechsler Scores	Whitaker & Gordon Adjusted Scores	Lanza Adjusted Scores
VCI		75	75	75
Vocabulary	26	8	8	8
Similarities	15	8	8	8
Comprehension	5	1	1	1
PRI		69	69	69
Block Design	10	5	5	5
Picture Concepts	12	6	6	6
Matrix Reasoning	8	4	4	4
WMI		77	77	77
Digit Span	11	7	7	7
Letter-Number Sequencing	9	5	5	5
PSI		65	65	65
Coding	20	4	4	4
Symbol Search	7	3	3	3
FSIQ		65	65	65

Note. * Highlights a difference in score

Table 9

Child 5 Wechsler, Whitaker and Gordon Adjusted, and Lanza Adjusted Subtest Scores

Subtest/Index	Raw Scores	Wechsler Scores	Whitaker & Gordon Adjusted Scores	Lanza Adjusted Scores
VCI		83	83	83
Vocabulary	31	7	7	7
Similarities	19	8	8	8
Comprehension	18	6	6	6
PRI		77	77	77
Block Design	22	6	6	6
Picture Concepts	15	7	7	7
Matrix Reasoning	17	6	6	6
WMI		86	86	86
Digit Span	16	9	9	9
Letter-Number Sequencing	13	6	6	6
PSI		53	50*	53
Coding	24	2	2	2
Symbol Search	8	1	0*	1
FSIQ		67	64*	67

Note. * Highlights a difference in score

Table 10

Child 6 Wechsler, Whitaker and Gordon Adjusted, and Lanza Adjusted Subtest Scores

Subtest/Index	Raw Scores	Wechsler Scores	Whitaker & Gordon Adjusted Scores	Lanza Adjusted Scores
VCI		71	71	71
Vocabulary	13	4	4	4
Similarities	8	6	6	6
Comprehension	22	9	9	9
PRI		88	88	88
Block Design	22	9	9	9
Picture Concepts	13	8	8	8
Matrix Reasoning	12	7	7	7
WMI		77	77	77
Digit Span	11	7	7	7
Letter-Number Sequencing	8	5	5	5
PSI		62	59*	59*
Coding	4	1	0*	0*
Symbol Search	9	5	5	5
FSIQ		70	69*	69*

Note. * Highlights a difference in score

Table 11

Child 7 Wechsler, Whitaker and Gordon Adjusted, and Lanza Adjusted Subtest Scores

Subtest/Index	Raw Scores	Wechsler Scores	Whitaker & Gordon Adjusted Scores	Lanza Adjusted Scores
VCI		67	67	67
Vocabulary	15	5	5	5
Similarities	9	6	6	6
Comprehension	5	2	2	2
PRI		79	79	79
Block Design	14	7	7	7
Picture Concepts	10	6	6	6
Matrix Reasoning	13	7	7	7
WMI		68	68	68
Digit Span	9	5	5	5
Letter-Number Sequencing	6	4	4	4
PSI		68	68	68
Coding	7	1	1	1
Symbol Search	7	7	7	7
FSIQ		64	64	64

Note. * Highlights a difference in score

Appendix H: Table 8

Table 8
Means and Standard Deviations of Index and FSIQ Scores

Index	Wechsler	Whitaker & Gordon (2012) Adjusted	Lanza Adjusted
VCI *Mean (SD)*	72.1 (9.32)	71.3 (14.53)	72.1 (9.32)
PRI *Mean (SD)*	73.1 (9.91)	73.1 (9.91)	73.1 (9.91)
WMI *Mean (SD)*	70.1 (11.45)	69 (13.93)	68.7 (14.58)
PSI *Mean (SD)*	60.4 (9.71)	58.1 (11.84)	59.3 (10.69)
FSIQ *Mean (SD)*	62.3 (8.83)	60.9 (9.56)	61.3 (9.79)

Appendix I: Table 13

Table 13

Wechsler FSIQ Ranges at 90% and 95% Confidence Intervals and Adjusted FSIQ Scores

Protocol	Wechsler FSIQ	FSIQ Range (90% CI)	FSIQ Range (95% CI)	Whitaker & Gordon (2012) FSIQ	Lanza FSIQ
Child 1	63	60–68	59–69	60	60
Child 2	43	41–49	40–50	40	40
Child 3	64	61–69	60–70	64	64
Child 4	65	62–70	61–71	65	65
Child 5	67	61–72	63–73	64	67
Child 6	70	67–75	66–76	69	69
Child 7	64	61–69	60–70	64	64