

Special Considerations for Score Interpretation Following Educational Disruption

Note: The following text is excerpted from an early draft of the WIAT-4® Technical & Interpretive Manual (Breaux, 2020). In its timely development process, the WIAT-4 team created helpful guidance for score interpretation in various situations, including times of “educational disruption,” which includes school closures related to the COVID-19 pandemic. The guidance below can be applied to all achievement tests, not only the WIAT-4.

Specifically, when considering how to interpret 2020 fall and winter achievement testing after lengthy spring school closures and widely variable educational access, we recommend considering the approach below to score interpretation. Assessment professionals and teams should think about the test interpretation process in terms of choosing the best reference group(s) for a particular student’s or group of students’ needs as well as additional variables that may have uniquely affected the student or group of students. Of course, many students will have been affected with less curricular rigor or access than their pre-COVID experience. Clinical judgement and multiple measures, as ongoing and consistent features of best practice, wisely anchor these situations.

***This document may be shared between colleagues
and your organization’s administration and leaders as needed.***

More questions?

**Please visit www.pearsonassessments.com/telepractice
and click “Contact Us” to connect with our team.**

Comparing Grade-Based and Age-Based Scores

For examinees in grades PK–12, the decision to use age-based or grade-based scores requires careful consideration. The following guidelines are based in part on the recommendations of Breaux and Lichtenberger (2016) and Farrall (2012). Four primary factors must be considered: (1) How the age-based and grade-based norm groups differ, (2) the academic areas tested and how the WIAT-4 results will be used, (3) the demographic characteristics and educational history of the examinee being tested, and (4) summer testing.

1. Consider how the age-based and grade-based norm groups differ.

Scores based on age norms compare the examinee’s performance to peers of the same age; however, peers in the same age group can be in at least two different grades and have different amounts of curriculum exposure. For example, if using the norms for Ages 5:0–5:3, the norms will include same-age peers in PK and K. One important factor that contributes to variability in the norm samples at the younger ages is that age requirements for kindergarten entrance vary widely across states in the U.S. Deadlines for children to turn 5 years old range from July 31 to January 1 of their kindergarten year (Education Commission of the States, 2018). In addition, the time of testing and the examinee’s birthdate affect the interpretation of age norms. If an examinee is tested early in the fall trimester and the examinee’s age is average for grade, then the age-based score compares the examinee’s performance to a group of age-matched peers primarily in the same grade and in the grade level below. However, if the examinee is tested later in the trimester or year (after his or her birthday), then the age-based norms primarily include examinees in the same grade and one grade level above.

Grade norms compare the examinee’s performance to peers of the same grade and trimester, which means that curriculum exposure is more consistent. However, peers in the same grade group can vary in age, particularly in the upper grades. For example, the Grade 9 norms primarily include students ages 14 and 15 (with a very small number of students in ages 13 or 16, depending upon the trimester).

The primary advantage of age norms is that they can be compared with other age-based tests and help reflect the skills of examinees who are young or old for grade; however, age norms mix examinees with different amounts of curriculum exposure. The primary advantage of grade norms is that they reflect curriculum exposure by trimester for each grade level; however, grade norms are less representative of students who are young or old for grade or have an unusual educational history.

When interpreting age-based or grade-based scores, consider the characteristics of the norm group in relation to the examinee at the time of testing and the extent to which the examinee is being compared to peers who are younger or older, or who have been exposed to more or less of the curriculum.

A summary of considerations for interpreting grade-based and age-based norms is provided in Table 3.4.

Table 3.4 Considerations for Interpreting Grade-Based and Age-Based Norms

Grade-Based Norms	Age-Based Norms
<ul style="list-style-type: none">• Is testing near the beginning or end of a semester (norm group)?• Is the examinee being compared to same-grade peers who are mostly younger or older?• Does the examinee have an atypical educational history such as retention, skipping a grade, a gap or disruption in education, etc.?	<ul style="list-style-type: none">• Is the examinee near the beginning or end of their normative age group (the examinee just had a birthday or is about to have a birthday)?• Consider the time of testing in relation to the examinee’s birthdate. Is the examinee being compared to same-age peers who have been exposed to more or less of the curriculum?• Is the examinee’s performance in math a focus of the evaluation? Math is heavily dependent on grade, rather than age.

2. Consider the academic areas tested and how the results will be used. Always be consistent in reporting age-based or grade-based scores across all subtests and composites for an evaluation. However, for a focused evaluation, age-based or grade-based scores may be preferred based on the domain area assessed. Mathematics, in particular, depends heavily on grade-level placement and curriculum exposure. For this reason, grade-based scores tend to be preferred when evaluating mathematics. Other domains, such as oral language, reading, and written expression, are less dependent on grade-level placement to develop, although reading and written expression may be more grade-level dependent in the early grades. When evaluating these areas, age-based scores are sometimes preferred.

According to IDEA Sec. 300.309 (2004) and many state-level guidelines, the eligibility team may determine that a child has a specific learning disability if the child does not meet age or grade-level standards in one or more domain areas. Standards include the oral language and academic skills and abilities that students are expected to demonstrate at various points in time. Some skills develop primarily with age, whereas other skills depend more on the grade-level curriculum. However, the same type of score should be reported consistently across subtests and composites, whether the skills depend more on age or grade. Examiners must follow state and district requirements based on IDEA (2004). Clinical judgment should be used to decide whether to use grade-based or age-based norms and to interpret the results according to the

examinee's demographic characteristics, overall academic performance, and educational history. In some cases, it may be most informative to interpret both age- and grade-based norms to determine whether a student is underachieving according to age- or grade-level standards.

When comparing WIAT-4 results with age-based (e.g., cognitive ability, language) test results as part of a comprehensive evaluation, age-based norms are necessary so that all test results are being compared to a similar reference group.

3. Consider the demographic characteristics and educational history of the examinee. In particular, the following circumstances warrant careful consideration:

Young or old for grade: Students who have been retained or, conversely, academically promoted, tend to be outside the typical age range of their same-grade peers. Generally, the greater the difference between the student's age and the average (or typical) age at the student's grade level, the greater the difference will be between grade-based and age-based derived scores. For example, a student who is old for grade, such as a 12-year-old in Grade 5, may obtain misleadingly low scores when age-based norms are used. For example, an age-based score on Numerical Operations would compare the examinee to 12-year-olds in Grades 6 and 7 who have been taught higher-level math skills. However, using grade-based norms could potentially mask difficulties by comparing this student to younger peers. With these cautions in mind, evaluate both age-based and grade-based scores for students who are young or old for grade.

On the cusp of a different norm group: When an examinee moves from one semester to another, or moves from one age group to the next, the results can be quite different when comparing scores based on the previous norm group and the current norm group. Results must be interpreted more carefully when examinees have just transitioned, or are about to transition, between norm groups. The current semester or age should be used for reporting results; however, comparing the scores from both norm groups and evaluating the overlap in confidence intervals can provide useful interpretive information.

Significant disruption in education: Disruption in education can occur due to personal factors (e.g., illness, trauma) or environmental factors (e.g., school closure, natural disaster, pandemic). To evaluate the educational impact of the disrupting event, consider the student's level of academic achievement before and after the event, the length of the disruption, the quality of education received during the disruption, and any supplemental instruction the student has received since then. If the educational impact of the disrupting event appears to be significant, the following guidelines are suggested:

- Interpret both grade-based and age-based norms;
- When interpreting grade-based norms, the current trimester should be used for reporting results; however, estimate the impact of the loss of curriculum exposure by comparing the examinee's grade-based scores to grade-based scores from an earlier trimester or grade level (the comparison grade level will depend on the degree of educational impact); and
- When reporting age-based norms, consider whether the norms include age-matched peers in grades above or below the examinee's grade level.

4. Consider the implications of testing during the summer. Although there is a clear line of demarcation between spring and fall norms during the summer, the reality is more complicated. An absence of formal or informal academic practice during the summer causes many students to gradually lose skills after the end of school (more rapidly in some domains than in others). Summer assessments may be another situation in which it would be best to report both age-based and grade-based scores.

Special Considerations for Score Interpretation

Clinical judgment is needed to decide whether to use grade-based and/or age-based norms and to interpret the results according to the examinee's overall academic performance and educational history. In some cases, neither grade-based nor age-based norms provides an ideal reference group based on the examinee's unique demographic characteristics or educational background. In these circumstances, clinical judgment is required, and the best decision may be to interpret both types of scores according to the guidelines provided here.

Perhaps most importantly, use confidence intervals to describe the examinee's academic skill level as an estimated range. In some cases, a more accurate estimation of the examinee's skill level may be at the upper- or lower-end of the confidence interval based on the factors described in this section.

Table 3.5 provides a summary of key factors to consider when deciding to use grade-based norms and/or age-based norms. Report the same type of score consistently across all subtests and composites for an evaluation.

Table 3.5 Guidance for Using Grade-Based and/or Age-Based Norms

Evaluate Both Grade-Based and Age-Based Norms		
Use Grade-Based Norms		Use Age-Based Norms
<ul style="list-style-type: none">• To obtain certain WIAT-4 composite scores when an examinee's age is out of level (age-based norms are not provided)• For an evaluation that focuses on curriculum-based skills	<ul style="list-style-type: none">• For examinees in Grades K–1 in a state with relatively early (July–August) or late (October–January) cut offs for kindergarten entry• For examinees who are young or old for grade (e.g., held back one or more grades or received advanced grade placement)• For examinees who experienced a significant educational disruption	<ul style="list-style-type: none">• For examinees of any age when comparing WIAT-4 results with age-based (e.g., cognitive ability, language) test results• For adolescents who have graduated from high school and adults ages 18–50 (grade-based norms are not provided)

The WIAT-4 grade-based and age-based norms are of comparable quality in terms of reliability and validity. Both sets of norms are divided into shorter time intervals at earlier grades/younger ages when academic achievement generally progresses most rapidly. Grade-based norms are divided into trimesters (fall is a 4-month trimesters, winter is a 3-month trimesters, and spring is a 5-month trimesters). Age-based norms are divided into 4-month intervals between ages 4 and 12; 6-month intervals for age 13; 12-month intervals for ages 14, 15, and 16; a 3-year interval for ages 17–19; and multiyear norms for the 20–30 and 31–50 age groups. This means that for Grades PK–7/ages 4–13, the intervals of the grade-based and age-based norms are generally comparable. For Grades 8–12/ages 14–18, the intervals of the age-based norms are larger than the grade-based norms. However, monthly norm intervals are unnecessary in the middle school and high school grades/ages because skill development, as captured on nationally normed assessments, does not progress as rapidly as in the younger grades/ages. Both sets of norms are developed using the same methodology. The trend of growth curves observed in the grade-based and age-based normative samples guided the development of both norms.