

Assessing Adult Learners' Skills

Why Scale Scores Are Better than Grade Level Equivalents

Do you need to assess adult learners' skills and understand assessment results? Are you confused by the debate about whether these assessments should report results as "scale scores" or "grade level equivalents"? The truth is, grade level equivalents reference the wrong performance indicators, report performance on a less relevant scale for adults, and don't tell you what you think they do. The main reason some people prefer grade level equivalents to IRT scale scores is they think grade level equivalents are more comprehensible to a lay audience. But even this is not true. With minimal explanation, scale scores linked to basic skill competencies can provide more meaningful information than grade level equivalents.

Scale scores linked to basic skill competencies can provide more meaningful information than grade level equivalents.

Scale scores developed through item response theory (IRT) report an individual's proficiency as a score along a fixed metric scale, with each score representing fixed gradations of difficulty in a person's proficiency at competencies or skills. CASAS assessments use a scale from under 150 to over 260. Other IRT scales may use different ranges. Competencies or skills can be placed on this fixed scale with the easiest competencies on one end and the most difficult on the other. Individuals can be assessed and placed on the same scale, matching their skills and proficiencies to the

skills and competencies on the scale. Individuals with the same scores on an assessment have the same proficiency, regardless of when or where they were tested.

Grade level equivalents, on the other hand, report test results in terms of a year and month in a school grade. The skills represented by grade level equivalent scores are those a child in that grade and month of school would have. These skills would vary with the skills of the children tested and used as the base of comparison, or the "norm group."

IRT scale scores are based on ability alone, while grade level equivalent scores are based on assumptions about knowledge, ability, and skill development. And where grade level equivalent scales are gross estimates of probable performance, IRT scales are also more refined and less biased.

Grade Level Equivalents Reference the Wrong Performance Indicators

Tests that report results as grade level equivalents usually measure the academic skills included in that grade level in elementary and secondary schools. While this is important, more critical for adults is whether they are able to *apply* these academic skills in *real world situations*. To provide useful information, assessments must measure whether adults have the reading and

Assessing Adult Learners' Skills

Why Scale Scores Are Better than Grade Level Equivalents

mathematics skills needed in the workplace and the community at large.

Grade level equivalent assessments measure where in the K-12 educational system an individual's skills lie. They try to tell you whether someone is performing in the same way a specific "norm group" performed, such as fifth graders in the fourth month of school (5.4), or students in some other grade and month. This type of information is not very relevant for adult learners. What is relevant for them is whether they have the knowledge, skills, and abilities essential for success in an employment setting and everyday life. The skills and competencies consistently required in a work setting are not the same skills that are included in the academic achievement tests.

CASAS' IRT scale scores provide a more appropriate reference to the skills and competencies adults need to function

Scale scores provide a more appropriate reference to the skills and competencies adults need to function effectively in the family, community, and workplace.

effectively in the family, community, and workplace. They assess individuals' reading and mathematics skills within the context of the tasks adults are likely to face. CASAS assessments, and the competencies they measure, were developed with input from business leaders, adult education program staff and learners, social service and

employment and training program staff, and other community organization representatives from across the nation. The assessments and competencies are regularly revalidated to reflect changes in adults' skill needs. They provide information that can help adult learners plan their future, adult educators effectively target instruction, and employers match adults to jobs.

Grade Level Equivalents Report Performance on a Less Relevant Scale

Grade level equivalent tests are "norm referenced." This means that they tell you how well people perform in relation to some comparison group, or "norm." Reporting adult assessments as grade level equivalents is inappropriate on two counts. First, it relates the adults' performance to that of other individuals, rather than to a defined set of skills. Second, it compares adults' performance to that of *children*, when skill expectations for the two groups -- one with over 16 years of life experience, the other with less than 16 -- are quite different.

Grade level equivalent scales compare test takers' performance to that of other individuals. These scales are not constant; they vary with the skills of the groups that serve as their bases of comparison. When we assess adults (and some would argue children as well), we want to know what skills they have and what they are able to do. We do not care how they compare to some chosen group

Assessing Adult Learners' Skills

Why Scale Scores Are Better than Grade Level Equivalents

of people. We only want to know whether they have the skills to do the job at hand.

Grade level equivalent scales are also built on the performance of children, and so they necessarily relate to the skills we look for in children. By focusing on these children's skills, they can easily ignore, and even hide, the real abilities and learning that adults have, while at the same time demeaning adults with the grade-month categorizations they use.

IRT scale scores compare test takers' performance to a defined set of skills. In developing its scale, CASAS first explored the range of performance of low-skilled adults, and put this performance on an unvarying scale. CASAS is then able to report the performance of the adults it assesses along this constant scale. CASAS, therefore, both focuses on *adult* skills and abilities, rather than those of children, *and* provides information about the set of skills and abilities an adult has, rather than how that adult's skills and abilities relate to those of other individuals.

Assessments whose accuracy is tied to a K-12 teaching structure cannot accurately capture adults' skills and abilities, or their learning gains.

Grade Level Equivalents Don't Tell You What You Think They Do

Assessments that report performance in terms of grade level equivalents provide a year and month score for each individual tested. That is, they say a person's skills match those of someone in, for example, the fifth month of the eighth grade. But this score doesn't represent some well-known, agreed-upon body of knowledge. What is taught in the fifth month of the eighth grade can vary between classes, schools, districts, cities, and states.

Because different grade level equivalent assessments are "normed" against different populations (different groups of individuals have served as their bases of comparison), a person could take two different tests on the same day and get two different scores -- even if the same publisher developed the two tests. You aren't able to tell what skills are represented by a particular grade level equivalent score, and you can't be sure that different tests will award the same score to the same set of skills.

There are other flaws in the development of grade level equivalent scales that make them suspect. Test publishers do not go out and assess the skills of individuals at each of the nine months of every school year in order to develop their scales. They usually just test a sample of students in the fall or spring, and *extrapolate* from these points to develop a

Assessing Adult Learners' Skills

Why Scale Scores Are Better than Grade Level Equivalents

continuum. This extrapolation assumes month-by-month, linear learning growth for the average student, which contradicts individual learning theory. For grade level equivalent assessments to be most accurate, they need to be administered within a couple of weeks of the time of year they were taken by the sample population that served as the norm. This is rarely possible when assessing *children*, let alone adults.

Assessments that are tied to grade level scores are designed to test individuals progressing through our nation's K-12 educational system. But adult education is not organized in this manner. Adults do not typically attend school six hours a day, five days a week, September through June. We do not expect adults to follow the same developmental progression as children, so we do not deliver instruction to them as we do to children. Assessments whose accuracy is tied to a K-12 teaching structure cannot accurately capture adults' skills and abilities, or their learning gains.

Scale Scores Are Just as Comprehensible as Grade Level Equivalents -- Only More So

The general criticism levied against scale scores is that they're "too abstract." They report a number that has no meaning outside the world of the psychometricians who developed them. Grade level equivalent scores, the story goes, are understood by *everyone*, because *everyone* went to school.

Grade level equivalent scores, however, are not understood by everyone. People may *think* they know what grade level equivalent scores capture, but they couldn't really define the skills associated with each score. They couldn't tell you what skills a third grader or eighth grader might have, even if there were an agreed upon standard within or across educational systems.

When business people -- potential employers -- see the CASAS scale scores and clear definitions of the skills mastered at each level along the scale, they have no difficulty finding within them the skills they need in their workplace.

IRT scale scores, on the other hand, can clearly define the basic skills along a fixed continuum of difficulty. The skill level descriptions developed by CASAS are easily comprehended and provide lay people with the information they need to gauge a person's skills. In fact, they provide much more information than is contained in the words "eighth grade, third month." Because scale scores represent clearly defined skills, they can be used to gauge individuals' skill levels in relation to their own educational goals, providing them with a clear indication of the skills they have yet to master in order to enroll in a job training program or succeed in an entry level job, for example.

Assessing Adult Learners' Skills

Why Scale Scores Are Better than Grade Level Equivalents

The business community is frequently quoted as saying it needs individuals with “eight grade” reading and math skills. This is not because business people know what is encompassed by the term “eighth grade skills,” or whether these skills match up with their skill needs. Rather, they say they need eight grade skills because that is the level defined as functionally literate, and it is the standard answer they have been giving for years. They resort to expressing their needs in grade level terms because they know of no other reference point or benchmark they can use.

In truth, employers do not care only about the academic skills grade level equivalent scores hope to reflect; they want to know whether adults can *apply these reading and mathematics skills to workplace tasks*. They want employees who can perform on the job. When business people -- potential employers -- see the CASAS scale scores and clear definitions of the skills mastered at each level along the scale, they have no difficulty finding within them the skills they need in their workplace. With scale scores, they know what they're getting; with grade level equivalents, they don't.

Conclusion

IRT scale scores referenced to competency descriptors tell you what you want to know in a clear and comprehensible manner. They measure and report what is important and relevant -- the skills adults need to function effectively on the job and in their communities. The single sheet of paper needed to translate scale scores into meaningful skill definitions would be easily understood by the American public, and, in only a short amount of time, readily accepted. After all, this is the same American public that has been reading about the rise and fall of SAT scores -- reported on the SAT's own 200 to 800 point scale - for decades.

Scale scores measure and report what is important and relevant -- the skills adults need to function effectively on the job and in their communities.