

### **THE ROLE OF ASSESSMENTS IN THE IDENTIFICATION OF GIFTED STUDENTS**

Assessments can be used for a variety of purposes, including identifying students for gifted programs; providing ongoing feedback to guide the instructional process; and to determine to what extent students have obtained intended goals (e.g., academic, affective) within a gifted program. The purpose of this position paper is to provide parents, teachers, and other advocates of gifted students with best practices endorsed by NAGC related to the first purpose--the role of assessments in identifying students for gifted programs.

NAGC believes that the process of identifying students for gifted and talented programs must be based on defensible measurement practices, including the process of selecting psychometrically sound assessments aligned with a program's goals and objectives; the administration and interpretation of the assessments by individuals with appropriate credentials or training; and the ethical application of decisions regarding gifted program placement. Further, NAGC believes that there are specific practices that are supportive of these measurement practices.

In recent years, there have been significant discussions regarding the role of traditional assessments in identifying students who are typically under-represented in gifted programs, including culturally and linguistically diverse and low-income gifted students, and the use of alternative assessments with these students such as nonverbal ability tests (Lohman, 2005). NAGC believes that assessments selected for use in the identification of gifted students must be sensitive to and appropriate for the characteristics of the students being assessed and must aim to be inclusive of students from different cultures, races, and economic circumstances. Program administrators should choose the most psychometrically sound assessments with appropriate norms for their population of students and programs and use them appropriately for selection (see Lohman, 2005). However, it is also imperative that test users and policymakers understand that alternative-type assessments are not panaceas to the issue of under-representation, each come with limitations in terms of reliability and validity, and that these types of assessments should never be used in isolation to identify gifted children.

Another issue that warrants consideration in the identification of gifted students is the decision to use group versus individual testing, which is often determined by the availability of resources and the characteristics of the children to be evaluated. More accurate assessment data may be obtained via one-on-one testing with very young children and children with special characteristics and needs such as those with dual exceptionalities. For these children it is important to have a tester who is sensitive to and experienced with the group being assessed as well as the training in the administration of the assessments.

NAGC believes that because the use of assessments is an integral part of the identification process, test users have a responsibility to ensure that all testing is conducted in a fair and ethical manner. Such practices include the appropriate storing of testing materials before, during, and after testing; training all personnel involved with the administration and/or scoring of assessments; utilizing assessments that are developmentally appropriate and for only the purposes for which they were developed; interpreting

assessment results to the appropriate audiences; and maintaining the confidentiality of students at all times.

While NAGC advocates for the use of multiple assessments in the identification of gifted students, NAGC also believes that combining disparate data from multiple assessments must be done in such a way as to identify not only those students who are in immediate need of instruction beyond the regular curriculum, but also those students who display the potential for high-level learning beyond the regular curriculum.

In order to best implement defensible assessment practices for the purposes of gifted program identification, NAGC supports the collaboration of multiple stakeholders, including teachers, parents, and other advocates of gifted children, as well as general education administration at the district and state levels. This collaboration works to ensure that the application of defensible measurement practices results in the equitable and consistent use of assessments for the purposes of gifted program identification.

### *Research-Based Practices Regarding the Use of Assessments for Identification Purposes*

Regardless of the type of assessments used for identification or whether students are assessed in groups or individually, there are five non-negotiable practices in the use of assessments as identification tools. First, the choice of assessment tools must match the definition of giftedness that has been determined by the state, district, or school. The degree to which the assessment tool is aligned with the definition of giftedness is an important aspect of validity. Further, any assessments used in the identification process also should be aligned with the gifted program's goals and objectives and desired outcomes for students as a result of participation in the program (Feldhusen, Asher, & Hoover, 1984). Program administrators must carefully consider the program's goals and objectives as well as the aptitudes, achievement levels, and other characteristics of students (e.g., motivation, persistence, interest) needed for success in the program in order to select instruments that provide the most reliable and valid data regarding students' potential for success.

Second, identification of gifted and talented students should not be based on a single assessment. Rather, multiple pieces of evidence should be collected that measure different constructs and characteristics aligned to the gifted program's definition, goals, and objectives (Callahan, Tomlinson, & Pizzat, 1993), ideally including a variety of format types (e.g., paper-and-pencil; performance assessment). Multiple pieces of psychometrically sound data obtained from a variety of sources result in a more comprehensive and thus, more accurate picture of the student on which to base selection. For example, if trying to measure mathematical ability, appropriate choices might include a selected-response, domain-specific mathematics achievement test (e.g., a multiple-choice assessment) and a constructed-response assessment (e.g., performance assessment) where the student solves problems in an authentic context. However, when multiple assessments are used, it is important that the assessments provide different types of information as well as measure the construct, i.e. mathematical reasoning ability, differently. For example, although multiple pieces of information are being collected, administering assessments that follow the same response format may unfairly penalize some students while benefiting others. Program administrators should consider the use of a variety of format types when considering the specific assessments that will be used in an identification process and choose assessments sensitive to the inclusion of under-represented groups, culturally and linguistically diverse, and twice-exceptional students.

Third, the assessment conditions should mimic as closely as possible a natural setting in which the student can fully demonstrate his or her knowledge, skills, and abilities. The greater the unfamiliarity of the assessment setting, the greater the potential for undue negative influences on a student's performance (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999). For example, testing some of a district's second-grade students in a high school cafeteria on a given Saturday, while other second graders are administered the assessments within their classroom context, unfairly penalizes those students who are assessed outside their natural setting.

Fourth, school system personnel have the responsibility to be well-informed consumers regarding the technical documentation of each assessment used for identification (Joint Committee on Testing Practices, 2004). Assessment developers or publishers should include information on an instrument's psychometric properties (e.g., reliability and validity) and only assessments with adequate psychometric properties should be used in the identification of gifted students. In the absence of this information, responsible persons should determine an instrument's reliability and validity for diverse populations prior to using the instrument in an identification process.

Fifth, school system personnel have the responsibility to ensure that persons who administer and score assessments used for identification are appropriately trained and that placement decisions are driven by defensible data and not based on personal relationships, political associations, or parental pressure.

### *The Variety of Assessment Types*

Assessments differ on dimensions such as: the degree to which they are standardized (e.g., using large national samples versus local samples); the type of response format (e.g., producing a response as opposed to selecting a response from a predefined set); the ways in which the material is presented (e.g., paper-and-pencil, computerized, oral); and the content (e.g., mathematics) or constructs (e.g., creativity) being assessed. NAGC believes that regardless of the type of assessment, only assessments that provide psychometrically sound information on students, regardless of language, culture, gender, race, or socio-economic status, should be used. The following are three types of assessments often used in identifying students for participation in programs and services for gifted learners.

1. *Objective-type instruments:* These types of selected-response assessments used for identification purposes range from standardized, nationally normed paper-and-pencil or computerized tests to locally developed and normed tests, including most of the aptitude and achievement tests used in schools as well as IQ tests (see NAGC position paper; "Use of WISC-IV for Gifted Identification"). When using these types of assessments, users should be fully aware of the test's purposes and have evidence of sufficient reliability of the test scores. In addition, test users should use assessments that have a sufficient ceiling for measuring students' aptitudes or achievement, lack item bias, and have support for the validity of the types of decisions that will be made based on the results of the assessment (Joint Committee, 2004).

2. *Performance assessments:* Performance assessments, authentic assessments, and portfolios are constructed-response assessments that may be used in the identification process. These types of assessments directly measure the domain-specific construct of interest. Examples of performance assessments include open-ended or extended-response items. For example, students might be asked to present arguments for or against a particular position on an issue, write in response to a prompt, or

conduct and write a report of a scientific investigation. Portfolios are examples of another type of performance assessment in which students present their ‘best pieces’ highlighting the strengths of each piece or a ‘work in progress’ where students illustrate their improvement over time. When using these types of assessments, test users have the responsibility of ensuring that high-quality training procedures for scoring students’ responses or rating students’ work are in place in order to achieve a sufficient standard for exact rater agreement (Moon & Hughes, 2002). The acceptable standard for rater agreement is 80% exact agreement between two raters evaluating the same student response.

3. *Rating Scales, Interviews:* Classroom observations of students’ behaviors, collected by the use of rating scales designed to assess student characteristics or behaviors, and student interviews can provide useful supplemental data, particularly on students whose talents may not be evident on traditional aptitude or achievement tests. NAGC believes that the use of rating scales and interviews should play only a supplementary role in the identification process. Collecting these types of information is very difficult to do well because all individuals are affected by bias and prejudice, even if only at a subconscious level. If these types of data are collected, it is important that one recognize that different genders, cultures, races, ethnicities, and social classes have different ways of communicating which may impact an observer’s/interviewer’s perspective on what behaviors constitute giftedness. It is also essential to recognize one’s own views and predispositions relative to these differing subgroups of the population. To guard against the introduction of observer/interviewer bias into the identification process, educators should use structured tools with inclusive, but specific and clear, criteria to guide the data collection process (Oosterhof, 2003). Program administrators have the responsibility to ensure that individuals collecting these types of data have sufficient training in both the use of the instrument as well as the manifestation of giftedness in differing subgroups.

### *Implications for Practice*

Program administrators are responsible for ensuring that:

1. the identification process and the assessments used are aligned with the program’s definition of giftedness;
2. the process includes the use of multiple assessments that are combined in a reasoned way that is not biased against any particular subgroup of students (VanTassel-Baska, 2007);
3. the types of assessments used have sufficient psychometric evidence supportive of decisions about students’ readiness for gifted programming;
4. all individuals involved in the assessment process have sufficient training in the administration and use of the assessments;
5. they themselves are fully informed about best practices in the field of testing as well as the latest research regarding the identification of gifted students; and
6. there is a process in place whereby the identification process is periodically evaluated to ensure it is reflective of best practices in the identification of gifted students.

Approved October 2008

## **Annotated Bibliography**

**American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.**

This reference outlines the standards associated with the testing process. The reference provides the criteria for the evaluation of tests, testing practices, and the consequences of using tests and is applicable for those who develop tests; those who select or review tests; those who administer and score tests; and those who use the results from tests for decision-making purposes, among others.

**Callahan, C. M., Tomlinson, C. A., & Pizzat, P. M. (Eds.). (1993). *Contexts for promise: Noteworthy practices and innovations in the identification of gifted students*. Charlottesville: National Research Center on the Gifted and Talented, University of Virginia.**

This report is based on a National Research Center on the Gifted and Talented study that involved the collection and evaluation of instruments and processes used in identification processes from districts across the nation. The work includes many of the projects that have been funded by the Jacob K. Javits Gifted and Talented Students Education Act.

**Feldhusen, J. F., Asher, J. W., & Hoover, S. M. (1984). Problems in the identification of giftedness, talent, or ability. In J. S. Renzulli (Ed.) (2004), *Identification of students for gifted and talented programs* (pp. 79-85), Thousand Oaks, CA: Corwin Press.**

This article, first appearing in *Gifted Child Quarterly* in 1984, discusses five steps associated with the identification of students for gifted programs: defining of program goals and the types of students to be served; procedures for nomination; procedures for assessment; individual differentiation; and validation of the identification process.

**Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, DC: American Psychological Association.**

The reference is a collaborative effort among the American Counseling Association, the American Educational Research Association, the American Psychological Association, the American Speech-Language-Hearing Association, the National Association of School Psychologists, the National Association of Test Directors, and the National Council on Measurement in Education. The Code serves as a guide for individuals who are test developers as well as test users and focuses on the (1) development and selection of assessments; (2) administration and scoring of assessments; (3) reporting and interpreting assessment results; and (4) informing of test takers.

**Lane, S., & Stone, C. A. (2006). Performance assessment. In R. L. Brennan (Ed.), *Educational Measurement* (4<sup>th</sup> ed.), (pp. 387-432). National Council on Measurement in Education & American Council on Education. Westport, CT: Praeger Publishers.**

This chapter provides an overview on the design, use, and validity of performance assessments for large-scale educational testing.

**Lohman, D. F. (2005). The role of nonverbal ability tests in identifying academically gifted students: An aptitude perspective. *Gifted Child Quarterly*, 49, 111-138.**

This article presents an overview of the different types of nonverbal ability tests, from individually administered to group administered and how selecting students on the basis of these types of assessments alone excludes many students who would profit from gifted and talented programs and includes many students for whom gifted and talented programs would be an ill fit.

**Moon, T. R., & Hughes, K. R. (2002). Training and scoring issues involved in large-scale writing performance assessments. *Educational Measurement: Issues and Practice*, 21(2), 15-19.**

This study investigated the amount of error introduced into students' scores from constructed-response items based upon the type of training raters received as well as the type of scoring used for evaluating student responses.

**Oosterhof, A. (2003). *Developing and using classroom assessments*. Upper Saddle River, NJ: Pearson Education.**

This text provides a thorough and succinct discussion of the issues involved with using assessments in an educational environment, from development of to the uses of to the interpretation of various types of assessments as well as the issues that need to be taken into consideration when using each type of assessment.

**VanTassel-Baska, J. (2007). *Alternative assessments with gifted and talented students*. Waco, TX: Prufrock Press.**

This text provides an introductory guide to the methods used in educational settings for identifying gifted students as well as discussion for the need to identify students from under-represented populations for gifted and talented programs. It also focuses on ways to assess the learning of gifted student through alternative means.