UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

EXAMINING THE MEASUREMENT INVARIANCE OF THE TRANSITION ASSESSMENT AND GOAL GENERATOR ACROSS PERCENT OF TIME SPENT IN GENERAL EDUCATION

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

JODIE DIANE MARTIN
Norman, Oklahoma
2013

EXAMINING THE MEASUREMENT INVARIANCE OF THE TRANSITION ASSESSMENT AND GOAL GENERATOR ACROSS PERCENT OF TIME SPENT IN GENERAL EDUCATION

A DISSERTATION APPROVED FOR
THE DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

BY

Dr. James Martin, Chair

Dr. Maeghan Hennessey

Dr. Joyce Brandes

Dr. Jim Gardner

Dr. B. Mitchell Peck

ACKNOWLEDGEMENTS

Special thank you to the staff of the Zarrow Center for Learning Enrichment at the University of Oklahoma, particularly Donna Willis, Amber McConnell, Linda Gill, and Jennifer Burnes. Donna and Amber, for your management and persistence in collecting the data for *TAGG* development, Linda for your management of the Zarrow Center operations, and Jennifer for your management of the data set.

Thank you to my committee members, Dr. Maeghan Hennessey for consistently asking me to explain why our work is important, Dr. Joyce Brandes and Dr. Jim Gardner for encouraging me to make this dissertation meaningful to teachers, and to Dr. Mitchell Peck, who taught me two very important lessons—to always run the frequencies when you manipulate a data set, and that theory rather analysis than should guide research.

An enormous thank you to my advisor, Dr. James Martin, whose knowledge of Transition is wide, and who allows others to contribute to its depth. On a separate note, thank you to Dr. James Martin, my supervisor, who allowed me the flexibility to maintain the work-life balance I needed to get a dissertation done.

Thank you to my family and friends, who each provided unique support. Thank you to my father whose expectations are consistently one step beyond where I am and to Dr. Marsha Herron, who consistently provides me much needed personal perspective. Finally, thank you to Donna, Amanda, and Jamar, who were always willing to give me opportunities to explore and share who I am outside the office when my life was consumed with work.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

LIST OF FIGURES

ABSTRACT

Federal law mandates that educators use transition assessment data when they develop their students with disabilities' IEP transition plans. Transition assessments typically provide scores for educators to include in transition plans, but educators must develop annual transition goals from those scores on their own. The *Transition Assessment and Goal Generator* (*TAGG*) was designed to bridge the gap between transition assessment scores and the development of annual transition goals. The purpose of this study was to test the configuration and structure of the *TAGG* theoretical model to ensure that the goals generated by the *TAGG* were valid for students who spent varying portions of their school day in general education. The study also tested the configuration and structure of the *TAGG* to understand whether the goals it generates were suitable for students at different grade levels, having exposure to different transition education concepts, and in different disability categories. Multigroup confirmatory factor analysis was used to assess for measurement invariance of the *TAGG* items on the Professional, Family and Student versions of the *TAGG*. Configural equivalence was found for each version of the *TAGG* regardless of students' participation in general education, access to transition education, grade level, and disability category, meaning those variables do not need to be considered when goals are generated by the *TAGG*. Some metric invariance was observed on the student *TAGG* for students having had transition education and on the family *TAGG* for students in the 9th grade and having specific learning disabilities and intellectual disabilities, suggesting that students and families viewed some of the behaviors as less related to success than for other groups of students.

CHAPTER ONE

Introduction

One of the purposes of the Individuals with Disabilities Education Act (IDEA) is to engage educators in the process of preparing students for transitioning from high school to life beyond (2004). Educator engagement includes collecting data that contributes to the individual education plan (IEP) teams' understanding the abilities of students, taking the lead on developing education plans that will lead students toward their self-selected goals, and delegating coordinated activities. Essentially, educators are the hubs in a system that includes outside service providers, families, and students; and they are expected to develop transition plans that respond to the future hopes of students when coordinating the services necessary to attain postsecondary goals.

To facilitate educators with transition planning, IDEA (2004) introduced the formal transition plan to the IEP process. Educators are required to begin including a transition plan for each student with his IEP by age 16, but are not prohibited from including the plan earlier. Under the law, educators are expected to administer transition assessments for developing plans to address students' goals for employment, education, and independent living. The Division on Career Development and Transition of the Council for Exceptional Children (DCDT), the professional association dedicated to progressive transition practices, encourages educators to base transition plan goals on formal and informal assessments, and to extend transition planning to personal and social environments (Neubert & LeConte, 2013).

Under IDEA (2004), educators are expected to base the plans they develop on their students' strengths, needs, and interests, which presents a unique challenge—

unlike typical students, students with disabilities generally have limited communication skills and abilities, ranging from communication disorders that limit speech to cognitive limitations that inhibit the knowledge of strengths and limitations necessary for choosing suitable goals and advocating for their goals (Thoma, Rogan & Baker, 2001). In response to these challenges, transition assessments have been developed to promote students' self-understanding and educators' understanding of their students' strengths, needs, and progress toward their goals for life beyond high school.

Developing Quality Assessments

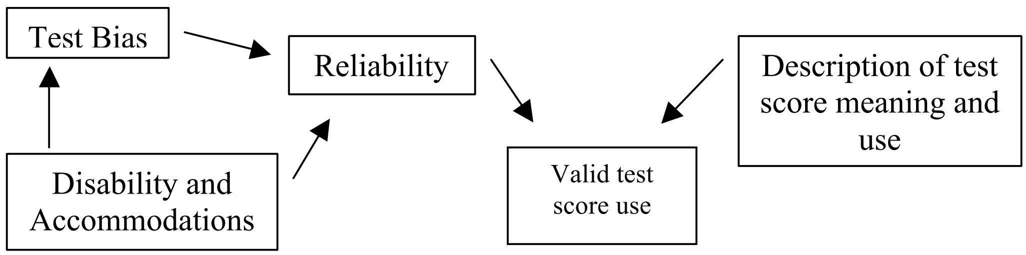
Transition assessment developers use *The Standards for Educational and Psychological Testing* as a framework for developing assessment and the technical manuals to be used with their tests (American Educational Research Association [AERA], American Psychological Association [APA], National Council on Measurement in Education [NCME], & Joint Committee on Standards for Educational & Psychological Testing, 1999). These standards are also commonly used among developers of academic and psychological assessments as well.

The Standards (AERA et al., 1999) outline the test properties developers should address to assure those who select tests that the tests have good validity and reliability. Among these properties are the empirical reliability of the test, promotion of valid use of test scores through a clear description of how test scores are developed and should be used, considerations for disabilities and accommodations, and an examination of test bias. This list presents these elements as discrete properties of a test, but when one of these properties is flawed, the tests do not work for measuring test takers accurately.

The relation of the test properties outlined in *The Standards* (AERA et al., 1999) can be understood using Figure 1.

Figure 1

Model for Understanding the Relations of Test Properties



In this model, test reliability and clear descriptions of a test scores' meaning and use both contribute to valid use of test scores. However, reliability is also impacted directly by undiscovered test bias, and directly and indirectly (through changes in test bias) by disability and changes made to tests to accommodate test takers with disabilities. Thus, this model proposes that as test bias and accommodations increase, the reliability of a test decreases, thereby reducing the validity of test scores. Furthermore, within this model, it does not matter the quality to which the test score interpretation sections are written, they do not improve the reliability of a test. Likewise, tests with exceptional reliability have limited validity when no clear description of score interpretation is provided.

Problems in each of the property areas in this model are produced from a range of sources. For example, problems with reliability (i.e., the consistency of the assessment across conditions such as time and raters) may stem directly from how items are written or understood, while issues associated with test bias are associated with factors outside the test, such as cultural experience, personal beliefs about what is being measured, or exposure to a concept. In this model, test bias impacts reliability by bringing in an unintended source of variation associated with the test taker. Disability and accommodations are observed factors that impact reliability in ways similar to test bias, but because they are rarely examined in the development of transition assessments, disability and accommodations may also be a source of test bias. Finally, issues related to descriptions of the meanings and uses of test scores stem from poor descriptions of how the scores were constructed, a poor quality scoring profile, and inadequate clarification of how and why test scores should be used.

Direct and Indirect Effects of Disability on Validity

Transition assessments are designed for use with students with disabilities, indicating consideration should be made for the impact of disability on their properties. Considerations are rarely made for the impact of disabilities and accommodations on test scores in general, and the selection of appropriate accommodations is typically left to the IEP team as part of IEP development. Although educators make good faith effort to select accommodations, they should not select accommodations that make tests easier unless the effects of accommodations are known. *The Standards* delegate responsibility for understanding the impact of accommodations to test developers.

Effects on test bias. Tests are designed to assess specific constructs, which are defined by the skills the test measures. Strong tests are perceived as measuring only the test taker's ability on the intended constructs. When tests have very low test bias against people with disabilities, it is assumed that a test taker with a disability would have the same score even if he did not have a disability. Driving road tests provide a powerful example of unbiased tests for people with physical disabilities. Although people with physical disabilities have impairments and use special tools to drive, they are not penalized for their impairment or for use of the tools. They are measured on the same set of skills, such as stopping at lights and using mirrors. In contrast, tests of physical ability that involve running or jumping are likely to be heavily biased against persons with physical disabilities who are likely to underperform specifically because they have a disability.

The inferences made about the test scores provided in these examples are based on the assumption that a test measures each test taker equally and fairly. On a test of physical ability, it seems intuitive that a person with a physical disability *should* score lower than a person without a physical disability, but unbiased tests should not rank test takers or generate scores that are a function of a disability, and tests should not be easier for persons without disabilities than for those with disabilities. Repairing test bias issues like these necessitates a reworking of a test's theoretical model and the test items, which will be discussed in a future section of this manuscript.

Effects on reliability. Accommodations are provided for students with disabilities as a workaround for the test bias associated with having a disability. For example, if a student has a disability that impacts his ability to complete math arithmetic, he may be provided a calculator to show that he is able to complete the steps of math problems that do not involve arithmetic. However, some accommodations—such as extended time on a time test—alter the administration procedures and perhaps the reliability of a test. Timed tests are designed to measure a test taker's ability on a trait during a timed condition. Providing extra time eliminates the condition, and changes the meaning of the test scores which, for a test of math fluency, is no longer the number of math problems a test taker can do in a certain time, but is now a test of the number of math problems a test taker can do all together.

Present Quality of Transition Assessments

The developers of transition assessments need to assure test users that their tests measure all of their target population reliably and consistently, and that the test scores are used in valid ways. Unfortunately, a review of the transition assessment literature found that disability and the use of accommodations was rarely considered as a factor in the reliability of test scores. Chapter Two of this manuscript examines this issue at length, but to preview, it is worth noting that of 52 assessments teachers reported using or that were suggested by the National Secondary Transition Technical Assistance Center (NSTTAC, 2013), only four considered disability-related variables.

Because disability and the use of accommodations have a powerful effect on the reliability of test scores, failure to consider their impact on transition assessments designed for students with disabilities has serious implications for whether the test scores are accurate. Educators are required by federal special education law to include transition assessment results in the development of transition plans, but doing so is moot if the scores are not meaningful and precise. At best, the scores are of low enough consequence that educators include them because they are required to do so, at worst, goals are developed using inaccurate test results and the student uses a year of his time working toward goals that do not match his actual needs.

Improving the Quality of Transition Assessments

Because disability and use of accommodations impact test reliability in different ways, their effects need to be evaluated using different methods. Group designs and differential item functioning (French & Finch, 2008) have been used to understand the effects of accommodations on academic tests. However, none of the commonly used transition assessments identified in Chapter Two of this manuscript undertook an empirical study of accommodations for their test, possibly because the test developers do not believe accommodations seriously threaten the reliability of their transition assessments.

When attention has been paid to the effects of disability and accommodation on test quality, the emphasis has been on intelligence and how test scores differ across groups. Using the model in Figure 1, however, an alterative and perhaps better approach involves understanding how disability impacts the reliability of a test, including an examination of the relations of the test items to one another, the strength of the factor structure, and the stability and consistency of test scores. Using plain terms and the example of a test of physical ability that will be given to students with and without disability, understanding the impact on disability may mean redefining the meaning of what good physical ability is and what poor physical ability is, and then finding ways (i.e. items) to measure those definitions. Once a revised and understood test has been evaluated and there is evidence that it is reliable for students with and without disabilities, educators may make valid use of the test scores.

Quantifying Disability

Each of the four transition assessments that considered the impact of disability on test scores classified students using a variable related to cognitive functioning—either IQ; predicted ability; or no disability, specific learning disability, or intellectual disability. It is acknowledged that intelligence is important to a test taker's ability to take a test. But, if a test is not measuring intelligence, accommodations may be used without concern that they will affect the reliability of the test. As Finch and French (2008) found, students who had an extended time accommodation on a timed test did not score any higher with the accommodation if they did not have the knowledge to answer the test items.

Instead, Vandenberg (2002) and the model in Figure 1 suggest that the disability variable of interest should somehow be related to the construct being measured. Because the category Students with Disabilities is broad and includes a range of students across of different ability levels and dimensions, selecting only one grouping variable is tenuous and does not necessarily capture the essence of how disability may impact a test's accuracy and precision. Ultimately, the disability-related variable selected for quantifying disability into the analysis of a test's reliability should represent an outside factor that is believed to directly impact a test. Referring back to the example of a test of physical ability, if the test examines rate of putting pegs in a board, meaningful variables that present bias would involve use of hands or fine motor skills, not simply having a physical disability.

An Applied Example

A new transition assessment has been developed to assess the nonacademic high school behaviors associated postsecondary employment and education among students with disabilities. The Transition Assessment and Goal Generator (*TAGG*; Martin, Hennessey, McConnell, Terry, & Willis, 2011) developers have followed *The Standards for Educational and Psychological Testing* (AERA et al., 1999) as a framework for ensuring their assessment is acceptable for use with students with disabilities who intend on entering postsecondary education and/or competitive employment. The reliability of all three versions of the *TAGG* has been assessed, and a scoring profile and description of the meaning of *TAGG* scores has been included in the *TAGG* technical/user manual.

One important feature of the *TAGG* (Martin et al., 2011) is the series of annual goal recommendations that are generated directly from each test taker's *TAGG* scores. This facet of the *TAGG* was designed to promote validity in the use of test scores. Because *TAGG* scores do not need to be interpreted by a third party educator, the developers have reduced the threat to the validity of test score use by reducing the need to rely on teachers to explain the theoretical concepts behind the *TAGG* to use the results for practical purposes. Educators are not expected to interpret *TAGG* scores and write transition plan goals, because annual transition goals based on each test taker's *TAGG* scores are provided directly to the test user as part of the scoring profile.

Presently, the *TAGG* (Martin et al., 2011) does not take into account differences between students when goals are suggested, which presents potential issues with the validity of the goals it suggests. Take, for example, two students—Raoul and Oscar—who both have disabilities and receive special education services. Raoul is categorized as having speech and language articulation impairment, and spends all but 15 minutes per week in general education classrooms. In contrast, Oscar has a mild intellectual disability and spends only about 20% of his school day in the general education classroom for art class.

Because both of these students have IEPs, they must have transition plans that include the results of transition assessments, along with goals that take into account the results of the assessments. Raoul and Oscar's transition specialist gives both of the students the *TAGG* (Martin et al., 2011) and both score 2 out of 9 on the Disability Awareness construct. The *TAGG* scoring profile suggested each of the young men have a goal for telling other people the supports and accommodations they need because of their disability.

To better understand the meaning of Raoul and Oscar's Disability Awareness scores, their transition specialist asked them each how they felt about having a disability and being in special education. Raoul told the specialist, "I don't think much about it. It's only for a few minutes and when I practice, I speak more clearly." Oscar replied by simply saying, "I'm just in small classes because I need them." The goal suggested by the *TAGG* (Martin et al., 2011) for the students may not be suitable for Raoul because his disability does not impact his general functioning enough to warrant talking to others about his disability. On the other hand, Oscar does not appear to have the language necessary to speak to others about his goal, and would likely benefit from the goal associated with a lower score—talking about his disability in a way that will get him the most help.

Disability group variable. Several considerations have been made for how disability should be quantified for this phase of *TAGG* (Martin et al., 2011) development. Chapter Two of this manuscript describes issues associated with the commonly used methods for grouping students with disabilities, along with the implications for using those methods. In large part, those methods are rejected because they do have a clear relation to the *TAGG* items in a way that might cause test bias. For example, using the frequency with which a disability occurs among students with disabilities does not identify a meaningful or useful source of test bias in an assessment of nonacademic behaviors, because it does not examine how the disability may impact the behaviors. There may be a relation between these types of variables, but to say that the frequency with which a disability exists causes a change in the reliability of an assessment's score is meaningless.

The skills and behaviors assessed by the *TAGG* (Martin et al., 2011) are learned and practiced in a range of settings. Some of the skills, such as leading IEP meetings, are most likely learned in special education settings, and students who spend little time in special education classes are less likely to have the skill *as it is measured on the TAGG*; therefore, a student may have and do the skill, but not the language to understand the item, a second test bias issue threatening the reliability of the *TAGG*. Reviewing the *TAGG* assessment items produced one disability-related variable that may have a direct threat on the test's reliability—the amount of time a student spends in general education.

In order to generate the most suitable goals for students, *TAGG* (Martin et al., 2011) scores must be precise and accurate. Ideally, the goals generated from the assessment will be included in a student's transition plan, and will be activities into which the student puts forth good effort and energy. Because IEP meetings are only held once each school year, if the goals are not suitable, or related to the student's needs, then substantial time is lost.

Using the model described in Figure 1, this study seeks to understand whether test bias is a product of disability when disability is measured by the percent of time a student spends in general education. The study will examine changes in the reliability of the *TAGG* (Martin et al., 2011) as a function of possible test bias. Because the goals generated by the *TAGG* function as the "valid use of test scores" component of the model and are generated directly from the *TAGG* items, this study will examine whether the items (and thus, the goals generated from them) are suitable for students regardless of their school placement.

Research Questions

This study seeks to understand whether the amount of time a student spends in general education variables need to be considered in the procedures for suggesting transition goals on the scoring profile of the *TAGG* (Martin et al., 2011). To answer this, the study will examine the theoretical model of the *TAGG* across the sample of *TAGG* students grouped by percent of time spent in general education. Therefore, this study seeks to answer the following questions:

1. To determine if configural invariance exists, are the items on the *TAGG* associated with the same constructs regardless of the percent of time students spend in general education?
2. To determine if metric invariance exists, do the items on the *TAGG* have the same strength of relations to their constructs regardless of the percent of time students spend in general education?

Implications

This study of the *TAGG* (Martin et al., 2011) has implications for the *TAGG* procedures for selecting goals, and transition assessment development methodology. The *TAGG* is an innovative assessment because it examines research-based *nonacademic* high school behaviors associated with postsecondary education and competitive employment, rather than behaviors typically valued by experts (e.g. researchers, employers, parents, postsecondary education administrators). This study takes the perspective that students need to participate in both general education and special education settings to learn the behaviors the *TAGG* associates with positive postsecondary outcomes.

***TAGG* procedures and development.** Although this is a study of the technical properties of the *TAGG* (Martin et al., 2011), the essence of this study is to understand whether the *TAGG* constructs have the same meanings for all students with disabilities and can be measured using the same behaviors regardless of the amount of time they spend in general education.

This study does not seek to understand whether population differences in actions associated with *TAGG*-identified nonacademic behaviors (Martin et al., 2011) exist based on the percent of time students spend in general education. Rather, the goal of this study is to understand whether the same set of observed behaviors are associated with the same constructs to the same degree regardless of the percent of time students spend in general or special education. If there are no differences, evidence will have been established that the *TAGG* is suitable for its target population, and that goals do not need to be further individualized based on the percent of time a student spends in general education. If differences do exist, the *TAGG* developers should adopt measures for accounting for these differences in their goal generation by weighting items accordingly and adjusting their algorithms for suggesting goals.

Transition assessment development. The *TAGG* (Martin et al., 2011) was developed through a review of the transition literature to develop a list of nonacademic behaviors associated with postsecondary education and competitive employment (McConnell et al., 2013). This study contributes to the understanding of the value of certain experiences for students with varying degrees of disability. For example, students whose disabilities do not significantly impact their functioning to a degree that they spend a large portion of their school day in special education may not need instruction in disability awareness. Conversely, students who spend time in general education may need extra practice generalizing the behaviors they learn in the special education classroom (Waldron & McLeskey, 2010; Zhang, 2011). Without an understanding of the relation of the amount of time students spend in general education and its impact on students' need to learn and practice the *TAGG* behaviors, educators risk pulling a student from general education participation for instruction when doing so does not provide meaningful benefit to the student.

Because each student's *TAGG* (Martin et al., 2011) scores are used to suggest goals, his goals are selected based on his ability within a construct. This process assumes that all students with disabilities need to learn the same sets of skills, a view that is held by most developers of transition assessments designed for students with disabilities. One potential outcome of this study is a method for quantifying the level of need created by a student's disability and then understanding the relation of his level of need with the behaviors used to define constructs. It is expected that because the *TAGG* behaviors are learned in different settings, students who participate in different settings will make different meanings of the *TAGG* constructs. On a broad scale, it would be of use to all developers of assessments for students with disabilities to consider the impact of differences in construct definitions caused by demographics related directly to their assessments.

CHAPTER TWO

Review of the Literature

This review of transition assessment literature has multiple parts. The review begins with a description of the transition assessments commonly used among educators of students with disabilities, and an analysis of the measurement-related validity and reliability evidence provided by the developers of those assessments. Specifically, the technical manuals will be reviewed for the criteria outlined in *The Standards for Educational and Psychological Testing* (AERA et al., 1999) and evaluated for the quality of the evidence.

It is worth noting here that all of the manuals excluded one key piece of validity and reliability evidence—an examination of their theoretical models and factors that may contribute to bias in their models (AERA et al., 1999; Vandenberg, 2002). This chapter includes a description of the multigroup confirmatory factor analysis (Cheung & Rensvold, 2002) procedures test developers may use to test their theoretical models, an expanded review of the use of these procedures in assessments that may be included as part of a battery of transition assessments, and traditionally used methods for grouping students with disabilities when group comparisons are attempted.

Finally, each student's *TAGG* (Martin et al., 2011) scoring profile includes a list of suggested goals based upon his scores on the assessment. However, the items included on the *TAGG* use language and behaviors students may only learn in certain settings, and thus, a student's *TAGG* scores may be a function of a factor outside of his ability—the amount of time he spends in general education and special education. This review will conclude with a description of traditionally used methods for grouping students with disabilities, the utility of these methods, and the implications associated with using traditional grouping methods. Based on the conclusions drawn from this section, rationale will be provided for examining level of disability as measured by the percent of time a student spends in general education.

Validity and Reliability of Transition Assessments

Several frameworks have been developed to promote appropriate use of educational assessments and evaluations. The frameworks range from theoretical models that seek to describe and construct validity (Cronbach, 1947; Messick, 1989), to professional ethical codes (AERA et al., 1999), to indices of acceptable levels of various properties of the tests (Cohen, 1992a; Salvia, Ysseldyke, & Bolt 2004), particularly where reliability is concerned.

To be clear, the definitions of validity and reliability are complex and multifaceted. For example, validity is generally included as a property of assessments, but the meaning of validity lies in the *accuracy* with which test scores are interpreted and used (Popham, 2008; Messick, 1989). Similarly, reliability is measured by reliability coefficients, complex mathematical calculations relegated to arbitrarily set levels of acceptability (Cortina, 1993), facing criticisms akin to those associated with significance criterion, effect sizes, and power in experimental research (Cohen, 1992b; Schmitt, 1996; Sijtsma, 2009).

Understanding the reliability of tests and the validity of test results requires sophisticated thought and an appreciation for the nuances of measurement and measurement theory (Messick, 1995). The purpose of this review is to expand on the model framework presented in Figure 1, and describe reliability and validity evidence educators should use when selecting transition assessments.

Because the purpose of this review is, in part, to discuss the validity and reliability evidence educators should examine when selecting transition assessments, attention will not be given to the full *Standards for Educational and Psychological Testing* (AERA et al., 1999), which will be described in full detail in the Validity section of this review. Rather, this review adopts the position stated in Section 3.2 of *The Standards,* which states that test users should select assessments based on the "publically available documented evidence of [the assessments'] technical quality…" (p. 4).

To address technical quality, *The Standards* (AERA et al., 1999) task test makers with providing evidence of the reliability and validity of the test, giving assurance that their tests are "as free as possible from bias," making the test accessible to people with disabilities through the use of modifications, warning users of the risks associated with misuse of the test, and providing score reports that contribute to and enhance understanding of the meaning of test scores. Although most test makers treat these components as separate features, this review takes the position that each of them contributes to the quality of the test, and thus the validity of the test results (Figure 1).

This review will describe evidence test makers can provide that their test meets each of these criteria and examine the quality of the evidence provided by makers of the commonly used transition assessments, including the *TAGG* (Martin et al., 2011). Table 1 lists the areas in which evidence should be provided and acceptable forms of evidence.

Table 1

Technical Evidence of Test Reliability and Validity

Facet	Evidence
Reliability	
Internal Consistency	Typically Cronbach's alpha
Homogeneity	Analysis of factor structure
Stability	Test-retest correlations
Inter-rater agreement	Correlations of scores, percent of agreement and disagreement
Parallel form	Correlation of scores from multiple forms
Test Bias	DIF Multigroup confirmatory factor analysis
Validity	
Scale Construction Scoring Profile Score Interpretation	Description of how scores are computed; sample scoring profile with instructions and description of each section; written examples for communicating meaning of scores

Next, the review will make suggestions specific to the *TAGG* regarding validity and reliability evidence to be included in its technical manual, particularly relating to potential validity concerns due to the variety of demographic differences within the *TAGG* sample. The review will conclude by examining the risks and ethical issues educators face when they use transition assessments that do not provide ample validity and reliability evidence.

Validity

The validity framework test makers adopted in developing their product shapes the selection of validity evidence included in their manuals. Ideally, test makers should adopt a framework that suits the purpose of their test. For example, if a test is developed to predict outcomes or assess content knowledge, the test developers should use a framework for validity that calls for the inclusion of evidence of the test's capacity to predict outcomes and appropriateness of content (Messick, 1995).

Traditionally, evidence of validity is broken into more than 12 types. Classical models of validity include content validity, construct validity, concurrent validity, and predictive validity (APA, AERA, & NCME, 1954). Criterion validity was adopted as a method for uniting predictive validity and concurrent validity (Cronbach & Meehl, 1955), though predictive and concurrent validity may still be addressed independently. For detailed descriptions of each of these types of validity, see any educational measurement textbook.

In 1989, Messick sought to redefine construct validity and presented a new framework steeped in philosophy and the nature of meaning, which makes it difficult to access by those outside academia (for example, his chapter in *Educational Measurement* included a table called *Taxonomy of Research Strategies from the Interaction of Inquiry Systems* to examine the inquiry systems developed by Leibniz, Locke, Kant, Hegel, and Singer [1989]), and many within.

Messick later published an easier to understand article that explicitly defined construct validity "not as a property of the test, or assessment…but rather of the meaning of the test scores," (1989, p. 741). Messick's revised definition described validity as having four unified facets—test score interpretation, use of test scores, evidence of the meaning of scores, and consequences of test scores—and six aspects—content, substantive, structural, generalizability, external, and consequential.

Furthermore, Messick's framework shifted the brunt (but not all) of responsibility from test makers, to test users.

The Standards for Educational and Psychological Testing used today draws from Messick's position that validity includes test properties that can be quantified—such as correlations with other tests for criterion validity, and test properties that cannot be quantified—such as the *expertness* of those who contributed to development of the test for content validity (AERA et al., 1999). *The Standards* do not prescribe the inclusion of validity evidence for each aspect of validity, because in practice, not every test needs to provide the same validity evidence (Messick, 1995). Tests that do not claim to predict likely do not need to provide evidence of the validity of their predictions, similarly, tests that claim to sample from a range of domains should provide strong evidence that each of the domains has been adequately sampled. Stated plainly, the evidence of validity test makers provide is adequate when it meets two criteria: (a) address the claims the test makers make about the test, (b) adequately answer questions test users have about the care taken to develop and purpose of the test.

Reliability

For those not comfortable with the constructivist, post positivism presented by Messick (1989, 1995) and the *Standards for Educational and Psychological Testing*, reliability presents an objective positivist alternative. Whereas the types of validity evidence developers need to provide could best be described as "provide enough to cover what needs to be covered," when reliability is addressed, there are specific areas in which evidence should always be included. Nevertheless, this does not mean that reliability is as simple as scanning that a handful of numbers meet a minimum criteria.

Reliability, like validity, is also multifaceted, and test developers need to demonstrate evidence of sound reliability theory in their materials.

Internal consistency. Internal consistency refers to the *strength of the relation between items* intending to measure a construct (Cortina, 1993). Internal consistency is typically measured using Cronbach's α[1], inter-item correlations, item-total correlations, or split-half reliability, all of which are, in essence, evidence that items are related. To clarify, none of these statistics provides evidence that a test is a measure of a unidimensional construct, though internal consistency is regularly supplied as evidence of unidimensionality. Cortina suggests using internal consistency statistics to confirm an identified single-construct structure after factor analytic procedures have been used to determine that the assessment is unidimensional.

Homogeneity. Homogeneity refers to the *relation of items to a construct*, which should be tested using factor analysis (Cortina, 1993) using either exploratory or confirmatory procedures. Evidence of these procedures includes factor eigenvalues, scree test analysis (Cattell, 1966), MAP test (Velicer, 1976), or parallel analysis (Horn, 1965), and factor loadings to assess the contribution of the latent trait to the scores on the items (Streiner, 2003).

Stability. Stability may also be called test-retest reliability, and refers to the consistency of scores across administrations when certain conditions are met (Taylor, 2009). Stability is generally assessed through the correlations of scores between

---

[1] A series of issues and concerns have been raised around the use of Cronbach's α as the measure of the relation between items and their constructs, ranging from its underestimation of the relation between items when it is calculated for a single administration (Cronbach, 1951; Sijtsma, 2009) to whether intercorrelations between items provide evidence of the relation of the items to a construct (Green, Lissitz; & Mulaik, 1977; Schmitt, 1996), particularly in multidimensional assessments that measure more than one construct (Cortina, 1993).

administrations using traditional conventions such the correlations of scores from one administration to the next.

Inter-rater agreement. Inter-rater agreement refers to the consistency of scores across observers in the same conditions. For example, if two educators observe a student at the same time and using the same tool, their scores should be very similar. Inter-rater agreement can be assessed many ways depending on the purpose and design of the assessment being developed (Saal, Downey, & Lahey, 1980). Inter-rater agreement should be included when human scoring may contribute to measurement error. No study included in this review followed this definition of inter-rater agreement, thereby omitting standardization procedures that would potentially reduce the variability and increase the reliability of test scores.

Parallel-form. Parallel-form reliability refers to the consistency of scores across different versions of the same measurement tool. Evidence of parallel-form reliability is generally assessed by the strength of the correlations between the scores on the forms using Cohen's conventions. Parallel-form reliability should be included when multiple forms have been developed with the intention of measuring the same constructs.

Test Bias and Measurement Invariance

The Standards (AERA et al., 1999) suggest that test makers provide evidence assuring test users that products are "as free as possible from bias…" (p. 4). The education community, as well as the general public, has a positivist fixation on The Fairness of the Test (Biesta, 2009), and assumes that tests measure all people in the same way, and that by assessing all people using the same test, we have controlled for all outside factors. This pattern of thinking precludes evaluation of test bias, and leads to inaccurate, potentially dangerous, value-laden, inferred group differences.

Test bias should be understood as any unintended demographic factor that contributes to test scores. When test bias is assessed, it is commonly done so to examine sociological demographics such as gender, socio-economic status, and ethnicity. When scores on a test are identified as being systematically related to membership in a group, test bias has occurred. Test users may have come to expect a certain degree of examination of demographic variables, but assessments of test bias should also include attempts to understand unintended factors that are unique to tests—such as exposure to the curriculum being measured (Vandenberg, 2002).

In the context of transition assessment, there are certain skills students may only have access to when they participate in special education classes. For instance, the core curriculum may no longer include explicit instruction in life skills. Thus, students who spend more time in special education are more likely to do well on an assessment of life skills than students who spend less time in special education. When this occurs, test scores are a function of placement, *not* a function of life skills knowledge.

Test bias can be assessed for various aspects of a test. For example, the language of a test can be evaluated for its accessibility to nonnative speakers of a language. Similarly, test content may be assessed for gender bias to understand whether both male and female students have the background knowledge or experiences necessary to answer items correctly.

Measurement invariance and equivalence examines tests for bias in the meanings of the constructs across groups. For example, on the *TAGG* assessment

(Martin et al., 2011) of behaviors associated with postsecondary employment and education, one element of utilization of a Support Community is accessing social services. The research studies used to develop the *TAGG* found that if a student accesses social services, he has a high level of ability in using his Support Community, a behavior positively associated with postsecondary employment and education. However, some students with disabilities do not need to access social services to enter postsecondary employment or education. For these students, using a Support Community may involve contacting family friends for employment opportunities, or accessing a trust to pay for further education. This mismatch in construct definitions is indicative of test bias at the theoretical model level.

Multigroup factor analysis and differential item functioning are used to empirically examine measurement invariance associated with test bias in measurement models (French & Finch, 2008, Meade & Lautenschlager, 2004; Vandenberg & Lance, 2000). Potential measurement invariance is reported using changes in fit statistics, factor loadings, and error variances. Differential item functioning can be used to examine systematic differences in performance by subgroups at the item level (Swaminathan & Rogers, 1990; Zumbo, 1999), with potential bias reported in R^2 effect sizes. These topics will be addressed in detail in the next section of this manuscript.

Disability and Accommodations

The Standards (AERA et al., 1999) specifically states that steps should be taken to plan for the use of accommodations and assessment of persons with disabilities, a position also taken by the International Test Commission (2001). In practice, accommodations are provided to reduce the effects of disability on test scores, but the impact of accommodations is rarely considered by test developers, as was done by Cohen, Gregg, and Deng (2005). Likewise, test developers rarely consider changes to tests' theoretical models for subpopulations of test takers who have disabilities as was done by French and Finch (2008), likely due to the cost associated with collecting and analyzing the data.

Score Interpretation

The fundamental consideration of unified validity (Messick, 1995) is test score interpretation and use, a responsibility shared between test makers and test users (AERA et al., 1999). *The Standards* task test developers with thoroughly describing the procedures used to build a scoring system, and including score reports that contribute to the understanding of scores.

There are no industry guidelines to address what should be included in score reports, but the nature and purpose of a test can guide the inclusion of appropriate information. For example, criterion-referenced tests should include a statement of the content the test taker mastered, while norm-referenced tests should indicate a description of the test taker relative to his peers (Popham, 2008). Similarly, Flaugher (1978) and Messick (1989) advocate a statement of the intentions and boundaries of test scores, for example, a statement that the test is designed to assess achievement, and not capacity, and that inferences about the meanings of test scores should not be generalized beyond the settings in which the scores were obtained.

Need for Review of Transition Assessment Practices

Because transition planning is federally regulated, the National Secondary Transition Technical Assistance Center (NSTTAC) was established as a resource for state education agencies and educators to improve the transition planning process. NSTTAC reviewed and catalogued published transition assessments based on a theoretical definition of each of eight categories (NSTTAC, 2013). For example, intelligence tests are defined as those assessments developed to assess cognitive ability. However, NSTTAC does not describe the origin of these categories, and assumes that all assessments fall in one of these eight categories.

The National Post-School Outcome Center (NPSO) provides technical support related to measuring postsecondary outcomes among students with disabilities. In particular, the NPSO supports educators with appropriately using Indicator 14 to measure the effectiveness of their special education transition programs. In essence, state and federal agencies use Indicator 14 to track the postsecondary progress of students with disabilities (Falls & Unruh, 2010).

The agencies use Indicator 14 data as the measure of a school's transition program. If students are employed and enrolled in continuing education, they are perceived as successful, labeled as experiencing positive postsecondary outcomes, and the school is perceived as having a good transition program. Likewise, if students are not employed or enrolled in continuing education after graduating high school, they are labeled as experiencing negative postsecondary outcomes, and the school transition program is labeled as one that needs improvement.

Review of Commonly Used Transition Assessments

The purpose of this review, in contrast to the purpose of NSTTAC's review (2013), is to understand the categories of transition assessments educators *commonly use*. Stated plainly, rather than describe transition plan elements federal law mandates, assessment practices academics encourage, and plan components administrators require, this review seeks to understand applied transition practice as it is occurring in the field. To do so, a review of the literature was completed using EBSCO and the search terms *transition assessment, disability, special education; commonly used transition assessments, disability, special education;* and *popular transition assessments, disability, special education*. Each search resulted in very similar findings—lists of *suggested* transition practices, but no search produced any peer-reviewed research related to the assessments teachers are using. Therefore, this review used a data analysis method to identify transition assessments teachers commonly used to develop their transition plans.

Method. The list of transition assessments included in this pilot study was generated from a pre-existing data set collected as part of a national study to develop the *TAGG* (Martin et al., 2011). As part of the demographic information collected, the researchers asked educator participants to indicate the extent ("Little," "Some," "Often") to which they used each of the following formal transition assessments: *AAMR Adaptive Behavior Scales*, *AIR Self-Determination Scale* (*AIR*), *Career Decision Scale*, *Ansell-Casey Life Skills* (*Ansell-Casey*), *Enderle-Severson Transition Rating Scales*, *Self-Directed Search Forms*, *Transition Planning Inventory* (*TPI*), *Transition-to Work-Inventory* (*TWI*), and *The ARC's Self-Determination Scales* (*ARC*). Teachers were also asked to list any other transition assessments they used, and the extent to which they used them. Some educators reported using state or district websites. These could not be included, because the educators did not indicate which transition assessments they used from the websites. Assessments educators used were coded as 1, regardless of the extent to which they stated they used them.

Participants. All educators who participated in first three phases of *TAGG* (Martin et al., 2011) development were asked to respond to questions about their transition assessment practices. One hundred fifty-six educators from 30 states responded to an item asking them which transition assessments they used in their transition plan development practice. Table 2 includes the demographics of the teachers who responded to the question.

Table 2

Demographics for Identification of Commonly Used Transition Assessments

Characteristic (n = 156)	n	%
Gender		
Male	12	8%
Female	143	92
Highest Level of Education		
Bachelor's degree	18	12
Some Master's courses	33	22
Master's degree	84	55
Ed.S.	11	7
Some Ph.D. or Ed.D. courses	6	4
Ph.D. or Ed.D.	1	<1
Missing	3	<1
Ethnicity		
Caucasian	134	86
African American	17	11
Hispanic	7	4
Other Hispanic	4	3
Filipino	1	<1
Multiethnic	1	<1

Table 2

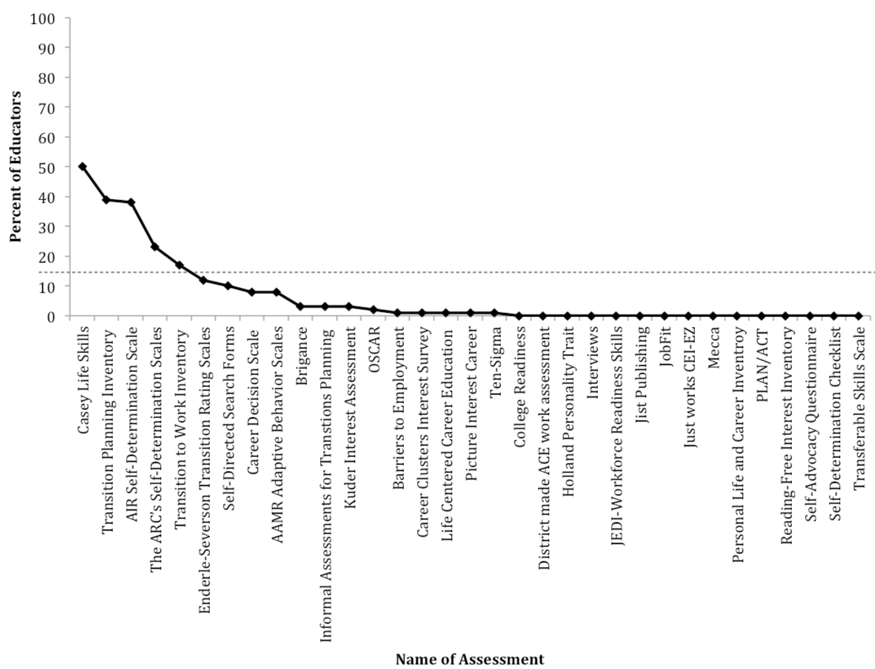
Demographics for Identification of Commonly Used Transition Assessments

Characteristic (n = 156)	n	%
Position		
Job Coach	7	4
Rehabilitation Counselor	4	3
School Counselor	4	3
Special Education Director	10	6
Special Education Teacher	117	75
Transition Specialist	38	24
Other	27	17
Number of Course of College Transition Courses		
0	63	40
1-3	61	39
4-6	23	15
7-9	9	6
Number of Hours of Transition Professional Development		
0-10	105	68
11-20	24	15
> 20	26	17
Grade Level Taught		
Below 9th Grade	26	17
9th Grade	104	67
10th Grade	118	76
11th Grade	133	86
12th Grade	136	88

Data analysis. Based on the educators' responses, a frequency count was used to graph the commonly used transition assessments. The line graph visually assessed for a change in slope, similar to reading a scree plot during an exploratory factor analysis (Cattell, 1966). A horizontal line was drawn to separate the assessments having wide usage from those used only occasionally. Figure 2 depicts the percent of teachers who use each of the transition assessments and the line that separates widely used transition assessments from occasionally used transition assessments.

Figure 2

Line Graph of Transition Assessments Teachers Reported Using



Thus, for the remainder of this review, the definition of *commonly used transition assessments* is defined as those transition assessments that are clearly used by more than a few teachers.

Once the list of commonly used transition assessments was developed, an inductive approach was applied to allow the categories and category definitions of transition assessments used by teachers to emerge. This approach promotes an understanding of transition assessment use as it is occurring in the field, rather than imposing a theoretical structure on educators' practice, thereby reducing the risk of value judgment associated with the rift evidenced by the vast literature on best practices in transition assessment and the dearth of research on educators' transition assessment practices.

The process of clustering the assessments began by examining the assessments to identify the constructs they intended to measure. This data was collected from the titles of the assessments, technical manuals, items, and literature-based descriptions of the assessments. For example, the *American Institute for Research (AIR) Self-Determination Scale*'s (Wolman, Campeau, Dubois, Mithaug, & Stolarski, 1994) title indicates it is intended to measure students' self-determination. The *AIR User Guide* (Wolman et al., 1994) describes the self-determination construct, and also clarifies that the assessment intends to measure capacity for self-determined behavior, and opportunities for acting in self-determined ways. Finally, the *AIR* occurs in the literature when researchers intend on understanding Self-Determination as a factor in various outcomes. Using the list of constructs identified during clustering of the assessments, inferences regarding the intended contribution of the assessment data to the transition plans were made.

Results. A total of 156 educators responded to a demographic data item asking them which transition assessments they used, and 63 of those educators wrote in assessments that were not included on the list. They reported using a total of 41 different transition assessments. Of those 41, eight were state transition websites that included transition education materials and assessments, and were excluded from the final list of transition assessments teachers reported using. Teachers used 33 different transition assessments to understand their students' future goals related to education, employment, independent living, and community involvement after high school to develop annual transition goals. Table 3 includes all of the assessments, and the percent of educators who stated they used them. The percent of educators who reported using each assessment ranged from 50% (n = 91) to 1%.

Table 3

Transition Assessments Commonly Used by Educators

Assessment	Percent of Educators	Number of Educators
Casey Life Skills	50	78
Transition Planning Inventory	39	61
AIR Self-Determination Scale	38	60
The ARC's Self-Determination Scales	23	36
Transition to Work Inventory	26	17
Enderle-Severson Transition Rating Scales	18	12
Self-Directed Search Forms	15	10
AAMR Adaptive Behavior Scales	8	13
Career Decision Scale	8	13
Brigance	3	4
Informal Assessments for Transitions Planning	3	4
Kuder Interest Assessment	3	4
OSCAR	2	3
Barriers to Employment	1	2
Career Clusters Interest Survey	1	2
Life Centered Career Education	1	2
Picture Interest Career	1	2
Ten-Sigma	1	2
College Readiness	1	1
District made ACE work assessment	1	1
Holland Personality Trait	1	1
Interviews	1	1
JEDI-Workforce Readiness Skills	1	1
Jist Publishing	1	1
JobFit	1	1
Just works CEI-EZ	1	1
Mecca	1	1
Personal Life and Career Inventory	1	1
PLAN/ACT	1	1
Reading-Free Interest Inventory	1	1
Self-Advocacy Questionnaire	1	1
Self-Determination Checklist	1	1
Transferable Skills Scale	1	1

Note. State websites removed

Overall, about 61% of the educators who participated used between one and three assessments, and another 22% of the educators reported they used four or more transition assessments. Based on Figure 2, the *Ansell-Casey* (Nollan, Horn, Downs, Pecora, & Bressani, 2002), *TPI* (Clark & Patton, 1997), and *AIR* (Wolman et al., 1994) can be clearly identified as commonly used by educators as part of transition planning. However, to expand the review, the *ARC* (Wehmeyer & Kelchner, 1995) and *TWI* (Liptak, 2012a) were added to the list of commonly used transition assessments. The remaining assessments were used by fewer than 15% of the respondents.

Five transition assessments appeared to be more commonly used than the other assessments: the *Ansell-Casey* (Nollan et al., 2002; used by 50% of educators), the *Transition Planning Inventory* (Clark & Patton, 1997; 39%), the *AIR* (Wolman et al., 1994; 38%), the *ARC* (Wehmeyer & Kelchner, 1995; 23%), and the *Transition-to-Work Inventory* (Liptak, 2012a; 17%). Of these all but one, the *TWI*, were included in NSTTAC's annotated bibliography that includes suggestions for completing transition assessments for students with disabilities.

Table 4 provides an overview of the validity and reliability evidence provided in the manuals for the transition assessments educators commonly use.

Table 4

Reliability and Validity Evidence Provided for Commonly Used Transition Assessments

	Reliability					Test Bias					
Assessment	Internal Consistency	Homogeneity	Stability	Inter-rater	Parallel Form	Model Fit	DIF	Disability	Scale Construction	Score Profile	Score Interpretation
Ansell-Casey				*	*					X	X
TPI	X		X	*	*			X		X	X
AIR	X	X	X	*	*						
ARC	X	X		*	*				X	X	X
TWI	X		X	*	*				X		X
TAGG	X	X	X	*	*	X			#	#	#

Note. * indicates evidence does not apply; # indicates unpublished evidence

Please note that Table 4 is simply a checklist and does not describe the quality of the evidence provided, which should be assessed in the context of an assessment's purpose as is done in this section. The evidence provided varied across the assessments, from the *Ansell-Casey* (Nollan et al., 2002) which provided no evidence of reliability, to the *TPI* (Clark & Patton, 1997), which covered nearly all of the areas the *Standards for Educational and Psychological Testing (AERA et al., 1999)* encourage developers to address.

Life skills assessments. Over half of the educators surveyed indicated they administered a life skills assessment to their students as part of transition plan development. All of them reported using the same life skills assessment—the *Ansell-Casey Life Skills* (Nollan et al., 2002). The skills assessed by the *Ansell-Casey* are most related to the independent living category of transition goals, and thus, it appears educators view independence as capacity for maintaining a home (housekeeping, cooking, financial planning), personal care (health, hygiene), interacting with others (professionalism, family relations), professional skills (finding a job, keeping appointments), future planning (education, training), and optimism.

Ansell-Casey Life Skills. The *Ansell-Casey* (Nollan et al., 2002) is a free, online resource designed to assess relations, work/study habits, planning/goal-setting, using community resources, daily living, budgeting/financial management, computer literacy, and connections with adults among youth in foster care. The website also endorses use of the assessment for youth in corrections, homeless shelters, and employment centers, as well as some special populations such as students who have alternative sexuality or who are from Native American families. The website does not endorse use of the assessment for special education related transition plan development, nor does it address disability-related issues such as use of accommodations, or disability awareness.

The *Ansell-Casey* (Nollan et al., 2002) can be taken online or in paper form, and can be completed by the youth only, the caregiver only, or both the youth and caregiver. Both assessments include 112 five-point Likert items in seven categories. The youth and caregiver forms include parallel items, with the items on the youth form reading, "I know…" and the items on the caregiver form reading, "Youth knows…" Users can also download a resource guide that includes goals related to areas in which the student has low scores.

A review of the literature related to use of the *Ansell-Casey* (Nollan et al., 2002) with students with disabilities produced two articles. The stated purpose of the first article was to develop a "life skills intervention with homeless individuals…" and included participants whose primary disability was related to mental health issues (Ihara, Wolf-Branigin, & White, 2012). The second article focused on using the *Ansell-Casey* as a measure of quality of life at the time of the assessment, and did not examine how using the assessment was related to attainment of goals (Helfrich & Fogg, 2007).

***Ansell-Casey validity and reliability*.** No technical specifications for the *Ansell-Casey* (Nollan et al., 2002) are available for immediate download. A manual for administration includes a chapter on reading the results, but does not discuss interpreting them. A search for technical information produced one study examining the reliability of scores when students who lived in one city took both the paper-based and web-based versions of the assessment. The study found no significant difference between the scores obtained using the web-based version versus the paper-based version (Bressani & Downs, 2002).

The *Ansell-Casey* (Nollan et al., 2002) is the most commonly used transition assessment among educators. The *Ansell-Casey Help Guide* (Casey Family Program [CFP], 2002) describes the assessment as, “A measure of youth confidence in their future and their permanent connection to caring adults” (p .5), which it measures using items focusing on personal care, general functioning, support, and optimism. The *Ansell-Casey Help Guide* does not provide a rationale for the items, or evidence that the items are related to constructs the *Ansell-Casey* claims to measure, but in one section of the manual, the “Looking Forward” subsection of the test is described as “a young person’s level of confidence in his success” (p. 35). In actuality, the *Ansell-Casey* samples items from the daily and life skills realm, that assess the independent living facet of transition, and includes the phrase “Life Skills,” in its name. To infer that it is an assessment of confidence is the type of “second-step” inference Popham (2008) refutes, and *The Standards for Educational and Psychological Testing* (AERA et al., 1999) discourages.

To promote valid use of the test scores, the *Ansell-Casey Help Guide* (CFP, 2002) includes sections on reading the youth and parent score reports. The instructions focus primarily on website use (such as how to see certain sections of the report), but some attention is given to how the score report is organized (by grouping items based on how the student rated himself). The manual references a resource guide available online from which goals can be selected for areas in which the student scored himself low. A second resource guide to interpret the results of the “Looking Forward” subsection is referenced. This guide does not enhance understanding of the test scores by contributing to the theoretical framework of the assessment, nor does it contribute to valid use of the test scores by making suggestions about how to use them.

With regard to the measurable properties of the *Ansell-Casey* (Nollan et al., 2002), the *Ansell-Casey Help Guide* (CFP, 2002) provides no reliability evidence to support the framework position that knowledge of life skills is a measure of confidence. The assessment claims to have been "designed to be as free as possible from…biases" (p. 5), but does not provide evidence that tests for bias were undertaken.

As a whole, the *Ansell-Casey Help Guide* (CFP, 2002) presents very serious violations of both the classical and unified conceptions of validity and reliability. If a label must be assigned, the *Ansell-Casey Help Guide*'s greatest violation may lay with *face validity* (Mosier, 1947). The assessment is beautiful to look at, computer-based, has point-and-click features, has a thick guide for using the website, has a resource guide with boxes and goals, and measures a construct educators seem to believe is important (life skills), but it provides no evidence that it is measuring what it claims to be measuring—either confidence, or life skills. The research community would benefit from developing an understanding of the factors that have led to the *Ansell-Casey* (Nollan et al., 2002) becoming the most commonly used transition assessment, particularly to improve our understanding of why educators select the assessments they select.

Self-Determination assessments. Many of the educators surveyed indicated they administered a self-determination assessment to their students as part of transition plan development. The *AIR Self-Determination Scale* (Wolman et al., 1994)—which frames self-determination as series of steps—assesses students' capacity and opportunity to follow those steps. In contrast, the *ARC's Self-Determination Scale* (Wehmeyer & Kelchner, 1995) was designed to measure students' present level of self-determination by measuring their self-determined actions.

Slightly more than half of the educators who used self-determination assessments indicated using both the *ARC* (Wehmeyer & Kelchner, 1995) and *AIR*, thereby indicating that educators seek to understand their students' levels of self-determination and whether their students are practicing self-determination skills. However, when only one self-determination assessment was used, 89% of educators reported using the *AIR*—which assesses ability and opportunity for self-determined action, while only 11% of educators reported using the *ARC*—which assesses level of self-determined action.

Self-determination, unlike the other constructs explored by transition assessments, is not directly connected with the three outcome areas defined by IDEA (2004)—education, employment, and independent living. However, based on their use of self-determination assessments, educators appear to view self-determination as a necessary component for attainment of postsecondary goals in those areas. Furthermore, because the model of transition planning emphasizes basing transition plan goals on assessment results, educators seem to believe that self-determination skills can be taught, or at a minimum, fostered.

AIR Self-Determination Scale. The *AIR Self-Determination Scale* (Wolman et al., 1994) was developed with four purposes: (a) assess the self-determination of students, (b) identify strengths and needs, (c) identify IEP goals, and (d) understand areas in which students need opportunities to practice self-determined behaviors. The *AIR User Guide* (Wolman et al., 1994) indicates that the *AIR* is appropriate for all school-aged children, however no norms are included to clarify differences in scores at various age levels.

The *AIR* (Wolman et al., 1994) Self-Determination Model has three broad components, with two steps in each component:

1. Thinking
    a. Identify and express needs, abilities, and interests
    b. Set goals to meet needs and interests
2. Doing
    a. Plan goals to meet needs and interests
    b. Act on goal plans
3. Adjusting
    a. Evaluate actions toward goals
    b. Alter plans if they did not work

The *AIR* structure assesses students' capacity for acting on each of the steps, and the opportunities they have in school and at home for acting on each step.

The *AIR* (Wolman et al., 1994) is available in three versions—student, educator, and family, each having corresponding items, with the items on the youth form reading, "I begin working…," and the items on the parent form reading, "My child begins work…" Items on the educator form are written conceptually, and include an example of the behavior. For example, for the student item, "I begin working on my plans to meet my goals as soon as possible," the corresponding educator item reads, "Student knows how to take actions to complete own plans successfully. Example: Kenneth knows how to follow through on a scheduled plan to complete his work accurately and on time."

All items on the *AIR* (Wolman et al., 1994) are scored using a five-point Likert-type scale ranging from Never to Always. Each domain the *AIR* assesses contains six items. The educator form of the *AIR* assesses five domains: Knowledge of Self-Determined Behaviors, Ability to Perform Self-Determined Behaviors, Perception of Knowledge and Ability to Perform Self-Determined Behaviors, Opportunity to Perform Self-Determined Behaviors at School, and Opportunity to Perform Self-Determined Behaviors at Home. The parent form assesses three domains: Things my Child Does, What Happens at Home, and What Happens at School. Finally, the student version assesses four domains: Things I Do, How I Feel, What Happens at School, What Happens at Home. All three versions also include open response items to understand the students' goals and his progress toward attaining his goals.

The *AIR User Guide* (Wolman et al., 1994) includes a short section on interpreting scores, and using scores, as well as construct definitions, and reliability and validity. To promote the inclusion of self-determination in transition planning, the *AIR User Guide* also includes a curriculum guide with activities that will potentially increase self-determination.

The *AIR* (Wolman et al., 1994) is widely used in research related to disability and self-determination, most often to observe the levels of self-determination across various groups of students, or to promote the inclusion of self-determination as a component of transition education (Carter, Lane, Peirson, & Glaeser, 2006; Carter, Owens, Trainor, Sun, & Sweden, 2009; Lee, Wehmeyer, Palmer, Soukup, & Little, 2008; Mithaug, Campeau, & Wolman, 2003; Lee, Wehmeyer, Palmer, Williams-Diehm, Davies, & Stock, 2012; Palmer & Wehmeyer, 2003; Wehmeyer, Field, Doren, Jones, & Mason, 2004). In contrast, little has been done to understand whether the model of self-determination assessed by the *AIR* is positively related to attainment of transition goals during high school, and postsecondary goals after high school.

*AIR validity and reliability***.** The *AIR* (Wolman et al., 1994) was developed out of academia to "assess and develop a profile of the student's level of self-determination," "identify areas of strength and areas needing improvement," "identify educational goals and objectives," and "develop strategies to build the student's capacities and opportunities to become more self-determined and prepare for a maximally independent adult life" (Wehmeyer, 1995, p. 9). Included in these purposes are the assumptions that becoming self-determined and acting in self-determined ways will lead to a certain level of independence, (though this should not be confused with the Independent Living construct of IDEA (2004), which refers to living in a minimally restrictive environment).

A good deal of the *AIR User Guide* (Wolman et al., 1994) introduction is spent describing the self-determination model, and providing a rationale for the structure of the *AIR* (Wolman et al., 1994). The items on the assessment first examine ability and then examine whether the student is allowed opportunities to use his ability. A blueprint of the framework is included as evidence that all domains of the theoretical model have been covered, a practice common in assessments of academic skills, such as end-of-course assessment technical manuals.

The score profile and interpretation chapters of the *AIR User Guide* (Wolman et al., 1994) do little to inform the test user about the meaning of the test scores. The profile gives a summed total of points for capacity and for opportunity, which are combined to give an overall percent of self-determination, but no indication is given as to whether this is a criterion-referenced score or a norm-referenced score. The *AIR User Guide* suggests using the scores to make comparisons across capacity and opportunity and forms, but educators are interested in doing more than comparing. Without more clarification, the test user must make very uninformed inferences and interpretations about what has been observed.

The *AIR* (Wolman et al., 1994) includes empirical evidence of the internal consistency, homogeneity, and stability of the assessment. A factor analysis was completed as evidence that the items were indeed measures of the unobserved construct, but no details are given to describe how the items were developed. The *AIR User Guide* (Wolman et al., 1994) attempts to provide evidence that no bias exists through a summary of mean comparisons across subgroups, but given that this assessment uses a highly multidimensional model, mean comparisons are wholly inadequate[2]. Finally, the *AIR* is designed specifically for students with disabilities, and mean comparisons were done between groups of students with mild disabilities and severe disabilities, yet the manual does not address the validity and reliability issues associated with the use of accommodations on the assessment.

---

[2] Flaugher (1971) outlines the risk of using group mean differences to assess for test bias, as well as the contribution of gender, overrepresentation, content sampling, and model selection play to test bias. Mean comparisons such as ANOVAs or OLS regressions assume that the instrument is technically sound and do not contribute to differences in group means (Berry, 1993).

The *AIR User Guide* (Wolman et al., 1994) is evidence of a good faith effort to develop an assessment with good technical properties without the manpower and experience large-scale publishers have. Clearly, Wolman et al. (1994) had reviewed other manuals and generated a list of the components a test manual should contain. Unfortunately, the gaffs in the manual are evidence of a misunderstanding of how tests function, and the ultimate purpose of measurement.

The ARC's Self-Determination Scale. The *ARC's Self-Determination Scale* (Wehmeyer & Kelchner, 1995) was developed to provide data as part of diagnosis and placement, evaluation of strengths and weakness, planning education and treatment, and evaluating the effectiveness of interventions. The *ARC's Procedural Guidelines* (Wehmeyer, 1995) clearly state that the scale was designed to enable students to evaluate their self-determination, strengths and limitations, and progress toward becoming more self-determined. The user manual also explicitly states that the *ARC* is best used to identify students' self-determination-related strengths and limitations, and that scores should not be used to compare students. Finally, it should be noted that the *ARC* refers to the Association of Retarded Citizens, which implies the *ARC* is normed on and targeted for use with individuals with cognitive impairments. Rather than through a survey of experts in the area of self-determination, the *ARC* was developed through a survey of individuals with cognitive disabilities about their beliefs about self-determination.

The *ARC* (Wehmeyer & Kelchner, 1995) model conceptualizes self-determination as an educational outcome, having four essential characteristics (Behavioral Autonomy, Self-Regulated Behavior, Acting in a Psychologically

Empowered Manner, and Self-Realization) and nine components (choice-making; decision-making; problem-solving; goal-setting and attainment; self-observation, evaluation, and reinforcement; internal locus of control; positive attributions of efficacy and outcome expectancy; self-awareness; and self-knowledge). Wehmeyer and Kelchner (1995) include in their framework the assumption that self-determination is positively related to quality of life.

The *ARC* (Wehmeyer & Kelchner, 1995) consists of one student form broken into four sections. The first section, Autonomy, includes 16 four-point Likert-type items ranging from, “I do not even if I have the chance,” to “I do every time I have the chance,” broken into two subsections. The second section, Self-Regulation, contains open-response cloze style items, and 16 four-point Likert-type items broken across three subsections. Section three, Psychological Empowerment, includes 16 forced-choice items, two cloze items, and open-response items to understand students’ future plans for living. Finally, section four, Self-Realization, consists of 15 agree/disagree items. The *ARC Procedural Guidelines* (Wehmeyer, 1995) includes a section on scoring the assessment and interpreting the profile, which does not include space for synthesis of the open-response items.

For research purposes, the *ARC Procedural Guidelines* (Wehmeyer, 1995) indicate the assessment is best used to examine program efficacy, environmental and individual contributions to self-determination, and to “evaluate the importance of self-determination on related outcomes and issues” (p. 96). Research patterns related to the *ARC* have followed those of the *AIR*. Studies have focused primarily on construct definitions and factors associated with high self-determination scores (Lachapelle et al., 2005; Lee, Wehmeyer, Palmer, Williams-Diehm, Davies, & Stock, 2012; McDougall, Evans, & Baldwin, 2010; McGuire & McDonnell, 2008; Shogren et al., 2008).

ARC validity and reliability. The *ARC's Self-Determination Scale* (Wehmeyer & Kelchner, 1995) was developed to provide assessment results test users and test takers can use to understand their level of self-determination, identify areas in which they need to improve their self-determination, and measure their progress toward becoming more self-determined. In contrast to the *AIR*, which relies heavily on student's beliefs about himself, the *ARC* purports to have operationalized self-determination. The *ARC Procedural Guidelines* specifically state that the *ARC* was designed for use in assessing strengths and needs in self-determination areas and for research purposes, and that it is not designed to diagnose students.

The *ARC's* (Wehmeyer & Kelchner, 1995) theoretical model describes self-determination as an educational outcome, or a set of behaviors that can be learned. The assessment is broken into five subsections that are described in the *ARC Procedural Guidelines* (Wehmeyer, 1995). Within each of those descriptions, Wehmeyer (1995) describes a series of behaviors associated with each of the subsections, but does not provide evidence that the behaviors are related to positive outcomes or improvement in self-determination. This direct connection between theory and selection of the items provides strong rationale for using the *ARC* for theory-testing purposes. To provide evidence that the *ARC* is suitable for educational purposes, a second chapter describes how items were developed by examining the behaviors of individuals who are self-determined versus those who are not.

The chapter dedicated to scoring and score interpretation focuses largely on interpreting and scoring open-ended *ARC* items (Wehmeyer & Kelchner, 1995). A short section at the end of the chapter describes how the scores can be used. Unlike the *Ansell-Casey* (Nolan et al, 2002), *TPI* (Clark & Patton, 1997), and *AIR* (Wolman et al., 1994), the *ARC Procedural Guidelines* (Wehmeyer, 1995) do not include a set of goals educators can use; rather, the *ARC Procedural Guidelines* simply encourage educators to review the scale, and write goals that address areas in which the student has scored low. The section also describes how the scores can be used for research purposes.

The *ARC Procedural Guidelines'* (Wehmeyer, 1995) chapter on the assessment's technical properties is crafted to address the key areas of validity and reliability thoroughly. Wehmeyer demonstrates a good understanding of validity and reliability evidence, and understands the meaning of the evidence he provides, but synthesizing all of the evidence would have improved the overall quality of the *ARC Procedural Guidelines*. The manual includes descriptions of factor analytic procedures to develop the constructs and select items and evidence of the internal consistency of the assessment, but is missing evidence of the test-retest stability of the assessment. Furthermore, no evidence or statement is provided to address the potential for test bias.

The *ARC Procedural Guidelines* (Wehmeyer, 1995) is several chapters long, and clearly, substantial time was dedicated to researching and writing the guide. The areas the procedural guidelines address—such as development of the items—are covered with exceptional depth and attention, and the areas that are not covered—test bias, stability—are fully ignored. The sections that are not included are not indispensable, and leaving them out fully may mean there are technical problems in those areas.

Career Assessments. The most common type of transition assessments educators used were related to career outcomes. These ranged from interest inventories to personality assessments to assessments of workplace readiness. Educators used one assessment far more than the others—the *Transition-to-Work Inventory* (Liptak, 2012a), which emphasizes identifying potential careers for the student. About 30% of the educators who used this assessment reported using another career interest assessment, indicating that educators view career assessments primarily as tools for identifying potential careers, not as measure of career preparedness.

Transition-to-Work Inventory. The *TWI* (Liptak, 2012a) is available for purchase and is designed to generate a list of career opportunities based upon the test taker's leisure interests. The assessment is targeted for use with adults, and does not target a specific population, or address disability-related areas such as communicating about a disability. The user manual states that the assessment is written at a junior high reading level, a method test makers employ to extend the assessment to those who have difficulty reading. A version of the *TWI* was developed for individuals with severe disabilities, but a review of that version noted problems with the scores not necessarily improving the administrator's understanding of the student (Prien & Hughes, 1993).

The fundamental assumption of the *TWI* (Liptak, 2012a) is that if a person enjoys an activity during his leisure time, he will likely enjoy a job that includes that activity. The assessment was developed by correlating a variety of leisure activities to the occupational clusters found in the *New Guide for Occupational Exploration*

(Ludden & Shatkin, 2001). Items were developed through a review of the literature on career counseling, and through consultations with vocational and rehabilitation counselors.

The *TWI* (Liptak, 2012a) consists of 95 five-point Likert-type items to be completed by the test taker. The *TWI* user manual includes a short section on item standardization, test-retest and coefficient alpha reliability, and construct and criterion validity. Liptak also included means and standard deviations for each of the career clusters for men and women. Beyond the user manual, no research was found related to use of the *TWI* among students with disabilities, or the predictive validity of the *TWI*.

***TWI validity and reliability*.** The *TWI* (Liptak, 2012a) was designed to be a brief assessment tool to match test takers' personal interests with employment or business opportunities. The *TWI Administrator's Guide* (Liptak, 2012b) provides a theoretical framework for the assessment, along with evidence of the need for the assessment. Because the framework for the *TWI* seems intuitive (that if test takers enjoy activities in their leisure time, they will likely enjoy activities in careers that also include those activities), the evidence for the framework is limited to one author, who is also the author of the test. Since the introduction includes many references to the importance of leisure counselors, a less savvy reader of this manual may misinterpret the quality framework presented.

The items included on the *TWI* (Liptak, 2012a) were selected to "develop an inventory that measures an individual's non-work interests" (p. 4). The items were developed through a review of the research and interviews with experts in the field, but details are included to explain the process of selecting items. To eliminate potential test bias, the *TWI Administrator's Guide* (Liptak, 2012b) asserts that the items were revised to remove references to gender, ethnicity, or other similar factors. Items were finalized through a process of correlations. This process is not described in detail, nor is a table with the correlations of the items with the career clusters included, which is key evidence of the homogeneity of the assessment.

A section of the *TWI Administrator's Guide* (Liptak, 2012b) on understanding *TWI* scores describes the meaning of each score range (low, medium, high), for which no rationale is given, and then refers test takers to the *New Guide for Occupational Exploration* (Ludden & Shatkin, 2001) for more information. The *TWI Administrator's Guide* clarifies that test takers may score high in multiple areas or in no areas. When this occurs, the guide recommends research begin with the area in which the test taker scored highest, somewhat undermining the need for labeling the score ranges. Because the *TWI* was not developed to be used specifically with IEP planning, no goals are linked to test takers' *TWI* (Liptak, 2012a) results, nor does the *TWI* include a scoring profile.

The *TWI Administrator's Guide* (Liptak, 2012b) section on reliability is brief and includes reliability and stability information for the *Leisure Search Inventory* (*LSI*, in Liptak, 2012b), upon which the *TWI* (Liptak, 2012a) was based, but does not include reliability or stability on the actual *TWI*. The section on validity includes correlations of the *LSI*, which in this section is referred to as the *LSI*/*TWI*, to another similar assessment. The guide also includes tables indicating the means and standard deviations for men and women in each of the career clusters, which the authors present as evidence that the assessment has good construct validity. Essentially, the guide takes the position that because women score high in career clusters traditionally associated with work women do—such as the arts and services, and score low in areas not associated with work women do—such as business, science, math, and engineering, the correct content was included on the assessment.

The likelihood that the *TWI* (Liptak, 2012a) will be used to make decisions that have serious consequences is low, and for the resources dedicated to its development, the *TWI Administrator's Guide* (Liptak, 2012b) is adequate for including a description of how the items were developed and evidence of the reliability of the constructs. The section on using the scores gives educators and test takers a direction for transition planning, but because of serious issues associated with the validity evidence of the *TWI*, its scores should be considered as only one very small piece of transition assessment.

Global Transition Preparedness. A handful of transition assessments have been developed to assess students' overall preparedness for transition from public education to life beyond. The second most commonly used assessment, the *TPI* (Clark & Patton, 1997), was intended to assess students in all areas associated with postsecondary outcomes, including self-determination, independent living, and employment. Educators who used the *TPI* sometimes used additional transition assessments, particularly those included in this review, but for the most part, the *TPI* was used without further formal assessment. The popularity of the *TPI* indicates educators view transition planning as constructed by knowledge of employment and careers (how to get a job and how to train for a job), independent living (health, home, and relations), self-determination, and community involvement.

Transition Planning Inventory. The *TPI* (Clark & Patton, 1997) is available for purchase and was designed to be used in conjunction with other transition assessments (Trainor, Patton, & Clark, 2005). It assesses nine constructs: employment, further education/training, daily living, leisure activities, community participation, health, self-determination, communication, and interpersonal relations. The *TPI* is targeted for use with students with disabilities, and the technical manual includes disability in the demographics of the sample along with ethnicity information.

The *TPI* (Clark & Patton, 1997) consists of three versions, school, family, and student. Each version contains 3-6 items per construct. The professional version is written at a conceptually more difficult level than the parent version. For example, for an item on knowledge about the job the student hopes to have, the parent version reads, “Knows about jobs in which he/she is interested,” and the professional version reads, “Knows the requirements and demands of his/her preferred occupations.” The student version also has open-ended items.

Clark and Patton (1997) explicitly state that they do not believe the *TPI* has the power to predict postsecondary outcomes. However, the decision to use the *TPI* indicates educators believe it will be of some benefit to their students. As has been the case with each of the assessments described in this review, no research exists that directly examines use of the *TPI* and postsecondary outcomes.

*TPI validity and reliability***.** The *TPI Administration and Resource Guide* (Clark & Patton, 1997) includes chapters on interpreting and using the assessment results, and the technical properties of the assessment. Clark and Patton (1997) take the position that transition assessments are designed to identify areas of need so that further assessments can be completed and/or goals can be developed. The guide also includes a list of goals educators can use in transition plans, but the goals are not tied directly to student scores.

The *TPI Administration and Resource Guide* (Clark & Patton, 1997) chapter on reliability and validity includes coefficient alphas as evidence of the strength of the items and constructs and test-retest reliability. Clark and Patton (1997) provide validity evidence through a description of how items were selected (content validity) and how other experts rate the content.

Unlike the *Ansell-Casey* (Nollan et al., 2002), which was developed by a nonprofit organization, the *TPI* (Clark & Patton, 1997) was published by a test development company and designed using the IDEA model of transition. The assessment includes subsections to address employment; postsecondary education; and independent living skills including daily living, leisure activities, community participation, health, and communication. The *TPI* also includes a subsection on self-determination, which, based on the common use of self-determination assessments, is an area educators also believe is valuable to assess as part of transition planning.

The *TPI Administration and Resource Guide* (Clark & Patton, 1997) describes the assessment's purpose as "…lead[ing] to meaningful planning and action to prepare students for dealing effectively with the demands of living in their communities after school is completed" (p.31). This model indicates employment and postsecondary education are facets of independent living, rather than separate activities. The structure of the assessment echoes that model by including employment and education as subsections rather than excluding them on what would otherwise be an assessment of independent living.

The *TPI Administration and Resource Guide* (Clark & Patton, 1997) includes a chapter on interpretation and score use. The chapter provides a set of guidelines for using the assessment scores, including instances when the scores should not be used (for example, if goals indicated by the scores conflict with students' family values); selecting the appropriate *TPI* (Clark & Patton, 1997) form to use; and when to complete further assessments. The chapter also contains the procedures for scoring the profile and interpreting the results for students with severe disabilities. Finally, the chapter includes a case study as a model for discussing and using *TPI* scores in an IEP transition plan.

A separate chapter in the *TPI Administration and Resource Guide* (Clark & Patton, 1997) addresses development and the technical properties of the test, with the technical properties included as a phase of development, rather than an afterthought. The chapter includes test-retest/stability correlations for each version (teacher, parent, and student) on each of the subsections. Cronbach's alpha is also included for each of the constructs, but is also provided as evidence of the homogeneity of the assessment, rather than the strength of the relations of the items. Missing from the chapter is a section to address reliability of the computer version and evidence that the *TPI* (Clark & Patton, 1997) was examined for test bias.

The *TPI Administration and Resource Guide* (Clark & Patton, 1997) addresses concepts in classical validity and provides evidence of content validity through a description of how the items were selected and how the format of the *TPI* (Clark & Patton, 1997) compares to similar assessments. Likewise, the evidence of criterion-related validity is addressed through a summary of the correlations of scores on the *TPI* with other assessments designed to assess the abilities of students with disabilities.

Of the manuals included in this review, the *TPI Administration and Resource Guide* (Clark & Patton, 1997) is the most technically sound. A framework for the assessment is provided, as is a sound rationale for the content on the assessment. Although there are a few problems with how the authors interpret the reliability evidence, of greater issue is the lack of evidence that the test properties are maintained across students of all ability levels, particularly given that there is a subsection on interpreting scores for individuals with severe disabilities.

Conclusion. Based on the commonly used transition assessments identified in this study, educators seem most interested in collecting data related to students' career interests, ability to care for themselves and live independently, and self-determination skills. Of these three areas, two can be tied directly to the stated purpose of transition assessment—promoting educator understanding of their students' progress toward their goals in postsecondary education, employment, and independent living. The *Ansell-Casey* (Nollan et al., 2002) provides educators with results that describe independent living skills, the *TWI* (Liptak, 2012a) with data reflecting students' career interests, and the *TPI* (Clark & Patton, 1997) with a broad, but superficial picture of students' preparedness for life after high school as defined by IDEA (2004).

That two of the five commonly used transition assessments are assessments of self-determination cannot be ignored. It is quite clear that educators value self-determination skills and believe they are associated with attainment of postsecondary employment, education, and independent living goals. Unfortunately, the assessments educators are using do not directly align with the purpose of special education as it is formally measured with Indicator 14. Stated plainly, there are no Indicator 14 items to assess students' postsecondary opportunities to make plans to meet goals, and feel good about having done so.

There is evidence that the behaviors associated with self-determination are related to positive postsecondary outcomes (Wehmeyer & Schwartz, 1997), but the present analysis was completed to understand the measures of the theory, not the theory itself, and the scores from the measures teachers are using to assess their students do not produce the data necessary to develop goals *directly associated* with postsecondary employment, education, and independent living goals. Use of self-determination assessments at the teacher level makes it clear that the mission of special education is very different for educators than for state and federal level officials. Ultimately, educators must develop transition plans that balance the demands of their state and local educational associations and the needs of their students.

Validity and reliability of commonly used transition assessments. Missing from each of the technical manuals of commonly used transition assessments is evidence that the assessments have comparable properties across the populations they are targeted to assess. Test developers and users have a responsibility to examine the comparability and fairness of tests (AERA et al., 1999; Messick, 1995), especially given the variety of characteristics among students who have disabilities. This issue will be addressed fully in the next section of this document.

The takeaway from this review should be that test properties are not limited to validity and reliability, nor can test properties be assessed using a checklist. This review was undertaken by a person who is knowledgeable about test development practices and understands test properties and their implications; yet, decoding the content and measurement jargon took considerable effort. That said, how can educators who have little to no training be expected to select transition assessments using the criteria described in *The Standards* (AERA et al., 1999), particularly when other professional organizations advocate selecting assessments based on what needs to be learned and the characteristics of the student (NSTTAC, 2013)? Furthermore, given the broad use of the *Ansell-Casey* (Nollan et al., 2002), should following the guidelines of *The Standards* prohibit educators from collecting data they deem necessary for developing quality transition plans?

This review did not address one facet of Messick's unified validity framework—the consequences of the decisions made using the test scores. Are the goals selected using the assessment results of serious consequence? It is the perspective of this writer that the potential for serious consequences increases when poor quality assessments are used *exclusively* for the development of transition plans, but that when plans are based on multiple assessments—formal and informal, with input from multiple observers, including the student (NSTTAC, 2013), and adjusted as frequently as necessary, risk can be reduced.

The Transition Assessment and Goal Generator

The *TAGG* (Martin et al., 2011) was developed to provide educators a way to assess the nonacademic behaviors associated with positive postsecondary outcomes and generate a set of transition goals related to those behaviors. *TAGG*-identified nonacademic behaviors include activities such as having a job during high school, learning from mistakes, and communicating about school performance. The purpose of the *TAGG* is to both assess students' ability in eight nonacademic construct areas and to generate annual transition goals that may be included on the students' IEP transition plan.

The *TAGG* (Martin et al., 2011) was designed for use with any high school student with a disability who intends on entering postsecondary education such as college or technical school, and/or postsecondary employment. Like most transition assessments, the *TAGG* was designed to be used as one piece in a battery of transition assessments. For instance, the *TAGG* does not assess students' career interests, nor does it address independent living skills such as home maintenance or financial literacy.

The fundamental assumption of the *TAGG* (Martin et al., 2011) is that action, not knowledge, will lead to success, and therefore, self-determination related to postsecondary employment and education is best assessed in the context of actions related to postsecondary employment and education. Each of the *TAGG* constructs includes three to six behavior items rated by how often the student acts (Rarely through Often). The educator version includes headings and definitions for each of the constructs, while the family and student versions do not.

Items for the *TAGG* (Martin et al., 2011) assessments were developed through a review of the nonacademic behaviors research identified as associated with postsecondary employment and education in the transition literature (McConnell et al., 2013). When a behavior was found to be positively associated with these positive postsecondary outcomes, it was considered for inclusion on the final assessment. During initial *TAGG* development, multiple versions were written for each item to ensure the items that best conveyed the development team's intended meaning were included on the final assessment. The development team was also responsible for determining the sequencing of the items, and as the experts in the field, assessed the difficulty of the items within each construct.

For most of the *TAGG* (Martin et al., 2011) constructs, the first few items assess knowledge about the behaviors, while the final items assess action on the skill. For example, in the construct Support Community, the first item relates to knowing who is a positive support and who is not, while the last items involves the student's action toward using supports such as employment or other social services agencies. Each version of the *TAGG* includes 31 items scored on a five-point scale for the educator and family versions, and a three-point scale for the student version, and three yes/no items.

At this stage of *TAGG* (Martin et al., 2011) development, each student's *TAGG* results are reported using a scoring profile that includes norm-referenced construct scores, a summary of present levels of performance, areas of strengths and needs, and suggested annual transition goals that are connected with the common core curriculum. For example, if a student scores low in the Disability Awareness construct, a suggested goal would be to write a paragraph explaining the ways his accommodations help him participate in school.

Each student's suggested goals are selected based on the student's measured ability within each of the *TAGG* (Martin et al., 2011) constructs. Sequencing the *TAGG* items to be conceptually more difficult was key to generating suggested goals. The procedure for selecting goals begins with using the construct score to identify the most difficult construct item the student mastered. It is then assumed that if the student mastered that item, he has also mastered the easier items within the construct (Guttman, 1950). It is also assumed that the student has not mastered any of the items that are more difficult than the item associated with their construct score.

Using this Guttman (1950) method of scale development and zone of proximal development (Valsiner & Van der Veer, 1999) to frame these procedures, it reasons that the student's goal should be the least difficult item the student has not mastered. (See Appendix A for a description of how *TAGG* (Hennessey, Martin, McConnell, Terry, Martin, & Willis, manuscript ; Martin et al., 2011) developers determined if an item was mastered.) For example, if *TAGG* (Martin et al., 2011) scores indicated a student has mastered the third item in the Goal Attainment construct—using a plan to attain goals—it is assumed that he has mastered the first two items—setting goals that match his strengths and needs, and creating short-term goals to attain long-term goals—and that he has not mastered the fourth, fifth and sixth items. Therefore, the student's *TAGG* scoring profile would recommend the fourth item—adjusting plans when they do not work—as a suggested transition goal.

This method for selecting goals neglects to consider the degree of impact of the specific *TAGG* (Martin et al., 2011) behaviors on postsecondary outcomes, and though theoretically sound, this method does not generate goals that are meaningful to individual students. Comparing the effect sizes of the behaviors on or correlations of the behaviors with postsecondary outcomes may resolve the issues associated with the sequencing of the items, but does not contribute to suggesting the most pertinent goals for individual students.

Devising goals that are unique to individual students presents a challenge. As the *TAGG* (Martin et al., 2011) stands, it does not take into account individual demographic variables necessary to generate goals that are unique to individuals. A good example of this can be found in the Employment construct. If a 14 year old, high school freshman has mastered the third construct item—having an unpaid job for a family member—the *TAGG* is designed to suggest the next item as a goal—the student will have a paid job. Unfortunately, though this is a logical next step, this goal is not appropriate for the student because 14 year old high school freshmen typically do not have paid jobs. In fact, it may be argued that having a paid job should not be considered part of the Employment construct definition for students who cannot legally have paid jobs.

An infinite combination of variables may be used to improve the process for selecting goals for the *TAGG* (Martin et al., 2011) scoring profiles. The development team collected demographic data on a range of topics including access to instruction on some of the *TAGG* constructs, such as whether the student has had instruction in how to lead his IEP meeting, or which courses the student took in special education and general education, and selection of a variable that will make test scores unique to test takers should be based on the *TAGG* content (Vandenberg, 2002). For example, taking gender into account when goals are being suggested does improve the *TAGG* profile's sensitivity for the differences between males and females, but consequently raises concerns about whether goals should be generated based on a student's gender.

These behaviors serve to define the construct, and as the *TAGG* (Martin et al., 2011) stands presently, it is assumed that all of the *TAGG* test takers also use these behaviors to define the constructs. Some of the *TAGG* constructs specifically address disability-specific issues that students are unlikely to learn in the general education setting (such as disability awareness, and IEP involvement), while others include skills students may improve through practice in the general education setting (utilization of supports, and interacting with others, for example). The range of settings in which the *TAGG* behaviors may be learned and practiced, and the variance within the *TAGG* sample in the percent of time spent in these settings provide a demographic variable to be included in the framework for suggesting goals included on students' *TAGG* score profiles.

*TAGG validity and reliability***.** The technical manual for the *TAGG* (Hennessey et al., 2013) is currently in the development process; therefore, this review will address the present version of the manual. The *TAGG* technical manual states that the purpose of the *TAGG* (Martin et al., 2011) is to facilitate "educators' practice of writing academic goals to a new set of nonacademic behaviors associated with postsecondary employment and education" (p. 1) The language used to describe the *TAGG* as an assessment of *nonacademic skills* suggests that it samples from what may traditionally be life skills, vocational, or self-determination content in order to assess preparedness for postsecondary employment and education. The theoretical framework is not presented as extensively as in the manuals for commonly used transition assessments, and thus, educators who are considering the *TAGG* may not understand what will be learned from the test scores without seeing the items first.

The *TAGG* (Martin et al., 2011) assessment is organized into eight construct sections, each of which is defined in the technical manual. The rationale and process for the development of each construct section is described in the technical manual (Hennessey et al., 2013), with connections made to research. The behaviors selected for the items are based on empirical evidence that demonstrating them is positively related to postsecondary outcomes. The technical manual also explains the rationale for student, educator, and family versions of the *TAGG*. The technical manual briefly describes for whom the assessment is best used, but this statement is tucked at the end of the introduction.

The *TAGG* technical manual (Hennessey et al., 2013) does not yet include instructions for administration or scoring procedures, but does contain a very brief chapter on score interpretation. One of the key features of the *TAGG* (Martin et al., 2011) is the scoring profile that presents test takers' scores on each of the subsections of the assessment, a summary of strengths and needs, and IEP transition plan goals specific to the test takers' *TAGG* results. Detailed instructions should be included for how the scoring profile can be used. Additionally, substantial effort will need to be dedicated to clarifying how scores are generated, since the *TAGG* uses a complex scoring algorithm instead of a summed score.

The *TAGG* technical manual (Hennessey et al., 2013) chapter on validity and reliability includes summaries of the homogeneity of the items and their relations to their constructs, the internal consistency of the assessment, and test-retest stability. Correlations between the versions are also provided, but because the observations are done in different settings, these are not suitable evidence for parallel-form reliability. To meet this facet of reliability, correlations between the paper and computer versions of the *TAGG* (Martin et al., 2011) should be included in the manual.

Completing the *TAGG* (Hennessey et al., 2013; Martin et al., 2011) project includes the development of a technical manual that addresses the areas outlined in the

Standards for Educational and Psychological Testing (AERA et al., 1999). At this time, no evidence is included to demonstrate that the *TAGG* does not exhibit test bias or measurement invariance across within group populations, which, given the range of disabilities represented in the sample, is imperative for generating valid test scores and profiles. Because the *TAGG* assesses students in multiple domains, and because students at various cognitive levels may have different opportunities to experience the *TAGG* behaviors, the *TAGG* developers should provide evidence that their assessment does not under- or overestimate the scores for the students who take it.

Multigroup Confirmatory Factor Analysis

Any student with a disability can qualify for reasonable accommodations under Section 504 of the Vocational Rehabilitation Act (1973), but to gain access to the intensive education services provided by IDEA (2004), students with disabilities must have a need for specialized instruction. The qualifier, “a need for specialized instruction,” is undefined, and does little to narrow the variety of ability and functioning levels among students who receive special education services. If the student is deemed eligible for services, he is assigned one of the 13 categorical labels described in Figure 3. It is these categories that make up the population of *Students with Disabilities*.

When test developers, researchers, and educators refer to Students with Disabilities, they mean all students who have some physical or mental condition that limits their independent functioning (Finkelstein, 2002), regardless of their IDEA categorical label. We develop theoretical frameworks and curricula for use with students who have disabilities (Kohler, 1996; Martin & Marshall, 1995; Wehmeyer, Agran, & Hughes, 1998), silently assuming that because a student has a disability, he will fit into our theoretical model, regardless of any other personal demographic factors.

A powerful example of this attitude and actions associated with it can be found in the full inclusion movement of the 1990s and early 2000s. It has always been the position of the special education community that students with disabilities should be educated in the most appropriate setting, but for a variety of reasons some members of the education community took the position that Students with Disabilities should participate in the general education classroom 100% of the school day (Kauffman & Hallahan, 1995). Subsequent research focused on why students with certain categorical labels were not best served in fully inclusive settings (Lewis, 1994; MacMillian, 1996; Mesibov & Shea, 1996), though it should be noted that even that research neglected the differences among students within the disability categories.

It is not the intention of this review to examine the Full Inclusion movement, but the lesson learned was largely that there is no one educational model or framework of inclusion that fits all students with disabilities, or even one category of students (Kaufmann & Hallahan, 1995). Students with disabilities also have traits and attributes such as ethnicity, gender, socio-economic status, and behavior that contribute to how they fit into the educational models selected by their schools. Furthermore, it is worth noting that members of the disability community view disability as a trait or characteristic similar to how they view their ethnicity, gender, and other personal characteristics, not as a feature that defines who they are (Finkelstein, 2002; Oliver, 2004). From their perspective, developing educational models for Students with Disabilities is akin to developing educational models that are expected to be effective for all girls, all students who are Asian, or all students who come from high or low socioeconomic status. There is simply too much dissimilarity within these groups to make broad assumptions about their characteristics.

The Standards for Educational and Psychological Testing (AERA et al., 1999) directly addresses the potential for insufficient attention to differences within populations in its list of test developer responsibilities. According to *The Standards,* tests should be "as free as possible from bias due to characteristics irrelevant to the construct being measured…" (p. 4). In other words, tests are considered biased if unintended factors contribute to the test scores, and unbiased if evidence exists that unintended factors do not contribute to test scores through the use of the procedures described in this review.

To contribute to understanding of the research community's understanding of *The Standards*, this review will explain the meaning of measurement invariance and equivalence and the multigroup confirmatory factor analysis (MGCFA) methods test developers can use to ensure educators that their assessments are unbiased and therefore suitable for the target population. The review will describe the reasons for and purpose of the MGCFA methods and threats to valid use of test scores associated with not undertaking these analyses.

Once the methods are described, transition assessments beyond those commonly used will be examined for suitable, adequate, and appropriate evaluation of test bias, including a synthesis of the implications associated with marketing tests for use with Students with Disabilities when, in fact, a test does not measure the population equivalently. Finally, this review will conclude with an analysis of the implications of the review findings and suggestions to the developers of the *TAGG* for including MGCFA procedures to accurately measure all students for whom the *TAGG* is targeted.

Factor Analysis Framework

Commonly used statistical methods are designed to understand the relation of a predictor variable to an outcome. But predictor variables, such as self-determination and confidence, must be measured before their relations to outcomes are understood. To measure an abstract characteristic, a test is developed using a theoretical framework that associates a set of observable behaviors with an unobservable latent trait, referred to as a construct. The statistical evaluation of the relation of the behaviors to the construct or constructs is done using exploratory factor analysis and then confirmatory factor analysis (CFA).

Unlike regression models in which a measured predictor variable is associated with an outcome variable, CFA is designed to test the theoretical model, (i.e., the relation of the behaviors and the constructs to which we believe they are related). The quality of the model is assessed for overall strength using goodness-of-fit statistics, and at the construct-behavior level using factor loadings. Once the overall model is deemed suitable and individual behaviors demonstrate an acceptable relation with the latent trait, a test has been developed. The behaviors with acceptable factor loadings become test items, and evidence can be reported that the theoretical model is homogenous and measures only the construct or constructs it claims to measure.

Multigroup factorial analysis (MGCFA) examines the fit of a theoretical model across the subpopulations within the population for whom an assessment is designed. A grouping variable is selected, typically gender or ethnicity, to understand whether the measurement model “fits” equally across the groups. If the model suitably fits the two groups, it is determined that measurement is invariant or equivalent. When a model does not fit the groups, the model is deemed inequivalent, and thus, bias may exist (Kline, 2011). Dimitrov (2010) clarifies that tests of MGCFA invariance are not tests of the generalizability of a model from one sample to another. Rather, he maintained that they are assessments of “whether the construct has the same *meaning* for each group” for which the assessment was designed (p. 124), and Vandenberg and Lance (2000), in their review of MGCFA literature, point out that MGCFA can and should be undertaken as an extension to an already existing confirmatory factor analysis.

This review of the MGCFA procedures includes only measurement invariance and inequivalence tests associated with test development and scoring (Dimitrov, 2010). Complementary tests were excluded from this review because they function to confirm the initial tests included in this study (Meade & Lautenschlager, 2004; Vandenberg & Lance, 2000) and do not contribute to understanding MGCFA and ME/I. Likewise, tests of structural invariance were excluded because they are designed to assess the suitability of comparing scores (Cheung & Rensvold, 2002; Dimitrov, 2010), which is a necessary step prior to making mean comparisons, but which does not contribute to the model-fit assessment nature of this review.

Configural invariance. Evidence of equal form or configural invariance tests for equality of the number of constructs and the items associated with each construct across multiple groups. When inequivalence occurs at this stage, it is likely that the latent trait the test developer intends to measure is not exhibited by all of the groups via the same behaviors (Riordan & Vandenberg, 1994). When this occurs, a unique set of items is necessary for measuring each group (Byrne, Shavelson, & Methuen, 1989).

Figure 3

Example of Configural Inequivalence Across Two Groups

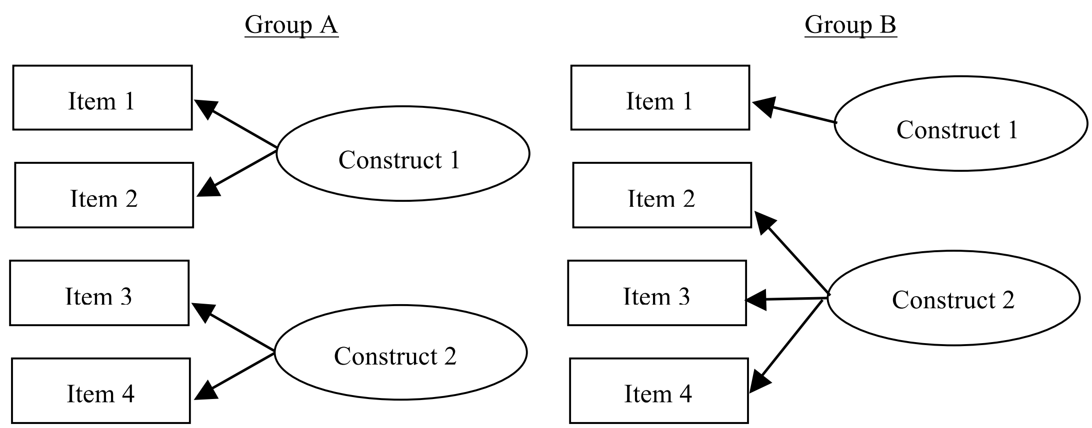


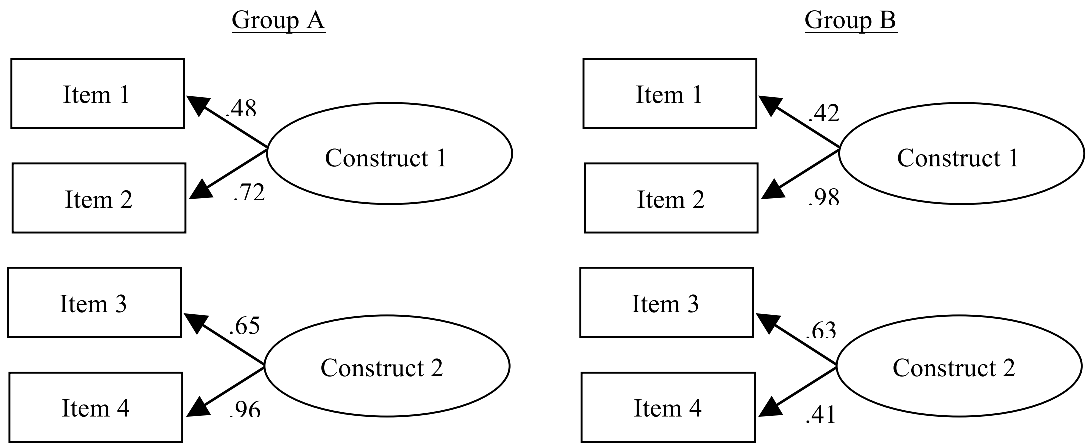
Figure 3 demonstrates a theoretical model in which items one and two load on construct one and items three and four load on construct two for Group A, while only item one loads on construct one, and items two, three and four load on construct two for Group B. This suggests the behaviors associated with the constructs are not equivalent across the two groups.

Metric invariance. If configural invariance is established, and the structure of the model (number of constructs and items associated with each factor) is suitable for each group, test developers should move forward to examine the potential for bias emerging from the relations of the items to their corresponding constructs by assessing for invariance of the factor loadings. Once evidence is established that the same items are associated with the same constructs to the same degree across groups, test developers can claim their test has *weak measurement invariance*.

Metric invariance tests examine the suitability of the model for each group and equality of the relations of the items to the constructs across groups, or, stated plainly, how different groups exhibit the latent trait a researcher is attempting to measure (Cheung & Rensvold, 2002). Figure 4 depicts a theoretical model having configural invariance for two groups, but no metric invariance.

Figure 4

Example of Metric Inequivalence Across Two Groups

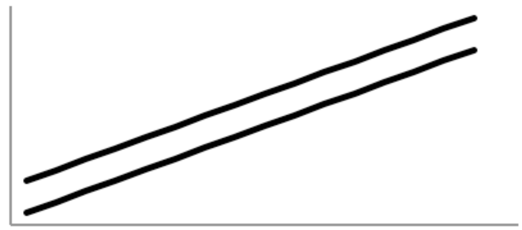


In this example, the contribution of the constructs to items scores means the constructs differ in how much they influence the item scores for each group. For items one and three, the contributions are about the same, and likely would not result in a noticeable difference in scores for the two groups. However, items two and four exhibit a large difference in contribution of the constructs to the item scores.

Issues in metric invariance can be further explained using regression terms. From this perspective, metric invariance indicates that the slopes of all of the groups are equal, and thus there are no extraneous factors contributing to the scores of one group or another. Figures 6 and 7 provide visual examples of these concepts.

Figure 5

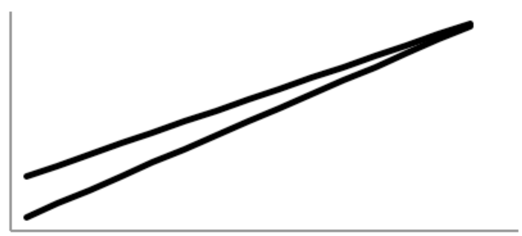
Regression-Modeled Example of Metric Invariance



Although the two groups in Figure 5 have different intercepts, both groups have equal unit increases on their item score for every single unit increase on the construct. This is evidence that the item has equal strength of relation to the construct for both groups.

Figure 6

Regression-Modeled Example of Metric Inequivalence



The two groups in Figure 6 have different intercepts, but also have unequal unit increases on the item score for every single unit increase on the construct. This is evidence that the item does not have equal strength of relation to the construct for each group.

Construct-level metric invariance. When configural invariance is established, and there is evidence that the same items are associated with the same factors for all groups, tests of the construct-level metric invariance can be used to assess the *overall strength of the relation between the items and the constructs* (Cheung & Rensvold, 2002). This is the assessment of the fit of the model to each of the groups, but it does not examine for differences in the relations of items to the constructs for each of the groups.

Item-level metric invariance. To understand the relations of items to constructs on a test, item-level metric invariance must be assessed. A review of the commonly used transition assessments found that some of the manuals made direct connections between the test items and recommended goals. However, to ensure the goals they recommend are suitable across groups, further examination of the strength of the relations between the behaviors and constructs needs to be completed. Once it is determined that both groups have equivalent factor loadings through tests of construct-level metric invariance, tests of item-level metric invariance may be completed to understand *strength of the relation between individual items and constructs* (Cheung & Rensvold, 2002).

When item-level metric invariance is established, test developers may suggest the same goals for each group based on the strength of the contribution of the behavior to the construct. When item-level invariance does not hold, test developers must examine the strength of the relations of the behaviors for each separate group in order to suggest the goals that have the strongest relation to the construct for each group.

Review of Multigroup Factor Analysis of Model Fit in Transition Assessments

We measure students to assign a number to their ability, but the numbers we assign them are only as accurate as the tools with which we take our measurements. The commonly used transition assessments reviewed in sections one and two of this manuscript provided varying degrees of assurance of the replicability, consistency, and stability of their measurements by including reliability estimates in their technical manuals. Missing from each of the manuals was an adequate examination of whether the theoretical models and test properties they provided were acceptable when disability was considered as a potential source of test bias (Figure 1).

Method

This review expanded the literature base to all assessments teachers reported using in the pilot study that identified the commonly used transition assessments, as well as the assessments recommended by NSTTAC (2013), less achievement and intelligence tests, and narrowed the focus of the search to the use of MGCFA. The initial list of assessments contained 93 tests, surveys, and inventories. In order to be included in this review, an assessment needed to have a technical manual available either online or through interlibrary loan that could be reviewed for evidence of MGCFA analysis of a disability disability-related variable.

The search for technical manuals was completed by searching for the name of the test, and the phrases “technical manual,” “user guide,” and “administrator’s manual,” in EBSCO and then in Google Scholar. If no items were returned, a second search was completed by searching the same search engines with the name of the test and “multigroup factor analysis,” and “measurement invariance.”

A total of 54 transition assessments met the inclusion criteria and were located from the original list of assessments used by educators or suggested for use as part of transition assessment. Table 5 contains a list of the transition assessments included in this expanded review and the number of MGCFA studies examining disability-related group differences.

Table 5

Number of MGCFA Studies Completed on Used and Suggested Transition Assessments

Name of Assessment	Number Studies Related to Disability	Subgroups
16 Personality Factors 5th Edition	0	
AAMR Adaptive Behavior Scales	0	
AIR Self-Determination Scale	0	
Armed Services Vocational Aptitude Battery	1	High, Medium, and Low by ASVAB Score
Becker Reading Free Interest Inventory	0	
Becker Work Adjustment Profile	0	
Bennett's Mechanical Comprehension Test	0	
Brigance Employability Skills Inventory	0	
Brigance Transition Skills Inventory	0	
Career Ability Placement Survey	0	
Career Beliefs Inventory	0	
Career Clusters Interest Survey	0	
Career Decision Scale	0	
Career Decision-Making System Revised	0	
Career Maturity Inventory	0	
Career Occupational Preference System	0	
Career Thoughts Inventory	0	
Casey Life Skills	0	
Choicemaker	0	
Differential Aptitude Test	0	
Enderle-Severson Transition Rating Scales	0	
Independent Living Scale	0	
Inventory for Client and Agency Planning	0	

Table 5

Number of MGCFA Studies Completed on Used and Suggested Transition Assessments, continued

Name of Assessment	Number Studies Related to Disability	Subgroups
Job Observation and Behavior Scale	0	
Jobs Search Attitude	0	
Career Exploration Inventory	0	
Kuder Interest Assessment	0	

Results

The initial list of assessments included in this review contained 93 tests, surveys, and inventories. Thirty-seven assessments were removed because no reference information could be found or because they were listed as assessments, but in fact were curricula. One study (Job Search Attitude Survey) was retained though references could not be located for its most recent version, and one study was retained because a website was found where it could be purchased (Ten-Sigma). A total of 54 transition assessments were located from the original list of assessments used by educators or suggested for use as part of transition assessment.

Two of the studies included in this review had been examined for model fit across groups selected using disability-related characteristics using MGCFA methods—the *Armed Services Vocational Aptitude Battery* (*ASVAB*), and the *Vineland Adaptive Behavior Scales*. The *ASVAB* (Ippel & Watson, 2008) study grouped a large sample of test takers by high general intelligence, average general intelligence, and low general intelligence. Ippel and Watson, who focused most of their effort on the high score group, found that although the factor structure and item loadings were equal for all three of the groups, the intercepts on certain subtests were not, thereby indicating some extraneous factor the test developers did not intend to measure was contributing to the scores of the participants in the high general intelligence group; Ippel and Watson call this "a serious violation of measurement invariance" (p.15).

The second study examined differences in the theoretical model of the *Vineland Adaptive Behavior Scale* (de Bildt, Kraijer, Sytema, & Minderaa, 2005; Sparrow, Balla, & Cicchetti, 1984) across children with mild mental retardation (MR), moderate MR, severe MR, and profound MR as part of an update to the technical properties of the *Vineland*. de Bildt et al. found that the *Vineland* factor structure was best suited for the mild and moderate groups, and that a different factor structure emerged for children with severe/profound MR. They attribute the differences in model fit to the theoretical model of the *Vineland* having been based on typical children, rather than children who are quite significantly different.

Discussion

This review examined a lengthy list of transition assessments to determine whether they had been examined for test bias, as per *The Standards for Educational and Psychological Testing* (AERA et al., 1999) and found that only two tests had undergone such analysis. It is worth noting that the two MGCFA studies on the transition assessments were completed by researchers seeking to understand the tests, not by the test developers.

Schmitt and Kuljamin (2008) point out that ignoring the issues addressed by MGCFA weakens the reliability of a test, which, according to Vandenberg and Lance (2000), is the essential evidence that a test is of good quality. The reasons the education community does not apply these methods more often are unknown, but could be related to time, sampling, inexperience, and lack of understanding.

This review, though its findings are limited, contributes useful knowledge to the disability community regarding the perspective of test developers. Clearly, the developers of these assessments did not see a need to examine them for test bias using sound methodology. A suitable next step is to understand whether test developers are looking for disability-related test bias using any other methods. If they are not, it can be inferred that test developers view disability as a static trait that impacts all people with disabilities equally. The next section of this manuscript will explore these issues.

In the context of disability and special education, it is fair to expect somewhat limited use of MGCFA methods due to sample size issues. MGCFA necessitates 150 participants (Meade & Lautenschlager, 2004). Collecting data on that scale, for a small population may not be feasible for an organization having limited resources or for groups where there simply are not enough participants available.

The two studies identified for this review provide evidence that theoretical models may, in fact, vary across subpopulations for which the test is intended. de Bildt et al.'s (2005) findings are particularly important to the special education community, because here we have an assessment designed for *Students with Disabilities*, yet the group for which it was designed was not as homogenous as the developers believed. Although a few of the assessments included in this review had been assessed for differences in measurement based on gender or English second language status, none of them had considered that type of disability could contribute to different manifestations of the theoretical model.

Suggestions for MGCFA and the *TAGG*

The information presented in this research review is intended to contribute to improving the technical quality of the *TAGG*, which recruited students in grades 9-12 with any disability and who planned to enter postsecondary education or competitive employment. The inclusion criteria produced a sample of students having IQs ranging from 40 to 126, and grade point averages of 0.00-4.00. The sample included students with autism, deaf-blindness, emotional disturbance, hearing impairments, intellectual disabilities, multiple disabilities, orthopedic impairments, other health impairments, specific learning disabilities, speech and language impairments, traumatic brain injuries, visual impairments, and other unspecified disabilities who lived in urban and rural settings. Students also differed on gender, socioeconomic status, ethnicity, and percent of time spent in the general education classroom.

The *TAGG* developers have followed test development practices established for the special education community. By including MGCFA analysis, the *TAGG* developers contribute to the special education research community by adopting a methodology that has strong utility given the range of characteristics among students with disabilities. The *TAGG*, on its own, is a unique assessment that contributes to the tools educators have for understanding the needs of their students. Including MGCFA procedures also enhances the statement educators can make about their students and the confidence they may place in the inferences they make about the test scores.

Finally, one of the long-term purposes of the *TAGG* was to develop an assessment of nonacademic behaviors associated with postsecondary employment and enrollment in further education. The data collected during *TAGG* development has also given the research team insight into the degree with which students are engaging in the *TAGG*-identified nonacademic behaviors. Before the researchers may move forward with analyzing what students are doing, they must be sure their measurement is true and accurate for all students who are being measured (Messick, 1995; Vandenberg & Lance, 2000). This is particularly the case given the methods the special education community uses to organize disability.

Traditionally Used Methods of Grouping Students with Disabilities

IDEA (2004) classifies students with disabilities into 13 broad categories. Every student who qualifies for special education services under IDEA must be assigned a categorical label, which is selected by the IEP team upon deeming the student eligible for special education services. Some categorical labels are selected based on evidence in medical evaluations—such as Speech Disorders, Other Health Impairment (OHI), or Autism, while others are based on the consensus of the IEP team—such as Specific Learning Disability (SLD), Intellectual Disability (ID), or Emotional Disturbance (ED). When students need services that are not covered under their primary disability, they may be given a secondary disability label that will entitle them to additional services.

Although the 13 IDEA (2004) disability categories embody a broad range of characteristics, the operational definitions of the categories present problems because they are not analogous and range from definitions of the disability to what are, essentially, "catch-all" categories. For example, ID is explicitly defined as a deficit in cognitive functioning with deficits in adaptive or functional behavior, and Orthopedic Disability is defined as a physical disability that impedes education performance. In contrast, ED is defined as a *learning impairment that cannot be explained* by cognitive, sensory, or health factors; and OHI is defined as a *health impairment that does not fit into any other categories*. Understandably, these catch-all categories were developed to eliminate the withholding of special education services to students whose educational performance could not be understood using the criteria for any other category, but given that many decisions are based on categorical labels, these catch-all categories do not guarantee their members' unique needs can be met.

IDEA (2004) disability definitions also present problems due to the shared characteristics *across* members of the categories and the rationales for the inclusion of certain categories. An easy to understand example involves the categories Deafness, Hearing Impairment, Deaf-Blindness, and Visual Impairment (Including Blindness). The Deafness and Hearing Impairment categories are defined based on the degree of the disability—Deafness indicates amplification does not contribute to improvement in educational performance, Hearing Impairment indicates the student has limited hearing, but is not deaf, which may entitle him to program-funded auditory devices. Based on these classifications, one would expect separate categories for Blindness and for Visual Impairment, but these do not exist. Finally, a third category, Deaf-Blindness, exists to ensure students who experience both are placed in programs that meet both needs.

The need for separate categories for Deafness, Hearing Impairment, Deaf-Blindness, and Visual Impairment (Including Blindness) is clearly driven by ensuring students have the most appropriate services. However, four other IDEA (2004) categories—SLD, ED, ID, and OHI—also share a good deal of overlap, largely because the operational definitions for the categories are overly inclusive (Hall, Peterson, Webster, Bolen & Brown, 1999; Semrud-Clikeman, Walkowiak, Wilkinson, Christopher, 2010), abstract or unclear (Stein, Klin, & Miller, 2004; Williams, Goldstein, Kojkowski, & Minshew, 2008), or influenced by the quality of data collected or the meaning IEP team members make of the data (Ysseldyke, Algozzine, Richey, & Graden; 1982; Ysseldyke, Algozzine, Rostollan, & Shinn, 1981). It should be noted that plenty has been written about the potential of non-disability factors contributing to placement decisions, as well as methodological issues associated with those factors (O'Connor & Fernandez, 2006).

The lack of homogeneity within IDEA categorical groups makes within group comparisons tenuous, while the overlap across groups make between group comparisons impractical. This review is designed to understand the solutions the developers of the commonly used transition assessments and NSTTAC recommended transition assessments adopted to group students when they examined the technical properties of their tests, including a discussion of the threats to measurement-related validity associated with their decisions. The discussion will also address other methods for grouping students with disabilities, methodological issues—such as sample size and theoretical soundness of decisions—and how these issues shape grouping decisions. Finally, based on the findings of the review, a solution will be proposed to the developers of the *TAGG* as they prepare to examine the equivalence of the *TAGG* model across various disabilities.

Review of Disability Groupings in Transition Assessments

This review examines the disability groupings used by the developers of commonly used and NSTTAC recommended transition assessments. The list of assessments included in this review was initially developed in the review of the

MGCFA methods used by developers of transition assessments. Because this review focuses on groupings and does not seek to assess the quality of the technical specifications of a test, manuals and reviews that used inappropriate methodologies to assess technical properties were retained as long as they contained a disability-related grouping.

Method

This review examines the disability groupings used by the developers of commonly used and NSTTAC recommended transition assessments. The list of assessments included in this review included all assessments teachers reported using in the pilot study that identified the commonly used transition assessments, as well as the assessments recommended by the National Secondary Transition Technical Institute (2013), less achievement and intelligence tests. In order to be included in this particular review, the technical manual or a review of the technical properties of the assessment first needed to be available via EBSCO, Interlibrary Loan, or Google Scholar. Second, the technical manual or test review needed to include disability-related groupings to address one or more technical properties of the test.

A search for technical manuals was completed by searching for the name of the test, and the phrases “technical manual,” “user guide,” and “administrator’s manual,” in EBSCO and then in Google Scholar. If no items were returned, a second search was completed by searching the same search engines with the name of the test and “technical properties.”

Retrieved technical information was examined for evidence that disability groupings had been used to understand the technical properties of the test. If disability was not addressed in the manual or among the technical properties of the test, the test was removed from the list of studies. If a test was developed using typical students, and those students were compared with a group described as “students with disabilities” or if a technical manual only referenced disability when describing the tests’ development sample, the test was removed from the list. When a test was retained, the groupings provided by the test developer or researcher were recorded in Table 6.

Table 6

Subgroups for MGCFA Studies on Used and Suggested Transition Assessments

	Number Studies Related to Disability	Subgroups
Life Centered Career Education Knowledge Battery	0	
McCarron-Dial Systems	0	
McGill Action Planning System	0	
Myers Briggs Type Indicator	0	
O*Net Ability Profiler	0	
Occupational Aptitude Survey and Interest Schedule	0	
Personal Capacities Questionnaire	0	
Prevocational Assessment Screen	0	
Scales of Independent Behavior-Revised	0	
Secretary's Commission on Achieving Necessary Skills	0	
Self-Determination Assessment Battery	0	
Self-Directed Search Forms	0	
Social and Prevocational Battery-Revised	0	
Supports Intensity Scale	0	
TEACCH Transition Assessment Profile	0	
Ten-Sigma	0	
The ARC's Self-Determination Scale	0	
Transition Behavior Scale	0	
Transition Planning Inventory	0	
Transition to Work Inventory	0	

Table 6

Subgroups for MGCFA Studies on Used and Suggested Transition Assessments, continued

	Number Studies Related to Disability	Subgroups
Vineland Adaptive Behavior Battery	1	Mild, Moderate, Severe/Profound by IQ
Vocational Cognitive Rating Scale	0	
Wide Range Interest-Opinion Test	0	
Work Adjustment Inventory	0	
Wiesen Test of Mechanical Aptitude	0	

The initial list of transition assessments included in this review contained 52 tests and inventories. Of those 52, the manuals of 45 could not be located, did not include any type of multigroup analysis related to disability, or had no examination of how the test measured students with disabilities differently. However, many of the manuals examined differences based on gender or ethnicity. Three manuals were requested via interlibrary loan but not received within the time frame of allotted for this review.

Table 7 includes the four of the 52 transition assessments educators reported using or recommended by NSTTAC that included evidence that the test had been examined for measuring students with various disabilities inconsistently.

Table 7

Disability-Related Groupings for Technical Properties of Transition Assessments

Test	Subgroups	Purpose	Findings
Armed Services Vocational Aptitude Battery	High, Medium, and Low by g	Model fit	Model did not hold for high-g group
Prevocational Assessment Screen	Mild, Moderate, Severe by ID	Group Mean Differences	
The ARC's Self-Determination Scale	No disability, SLD, ID	Group mean differences	
Vineland	Mild, Moderate, Severe/Profound by IQ	Model fit	Model did not hold for severe/profound group

The Prevocational Assessment Screen (Rosinek, 1984) and the *ARC* (Wehmeyer & Kelchner, 1995) each used group mean comparisons in their manuals, which are inappropriate because they assume no measurement difference (Vandenberg & Lance, 2000) on a test of measurement differences.

Independent studies were used to understand differences in the measurement properties of students at various ability (and disability) levels for two assessments—the *ASVAB* (Ippel & Watson, 2008), and the *Vineland Adaptive Behavior Scales* (de Bildt, Kraijer, Sytema, & Minderaa, 2005). These two studies focused on and found model fit differences across ability levels.

This review found that very few of the assessments used by educators and suggested for use by NSTTAC considered differences among students who had disabilities. When test developers examined the technical properties of their assessments for subgroups, they typically focused on age, gender, or race, and failed to address other variables that potentially contribute to differences in test scores, such as socio-economic status or national origin (AERA et al., 1999).

When tests included in this review examined disability groups, they sorted them based on some variable related to cognitive functioning—IQ or organizational psychology's analogous *g* (Sternberg & Hedlund, 2002), as depicted in Table 6. This grouping method correctly assumed that the theoretical models tested would not hold across cognitive ability groups, but does not address other ways disabilities may interact with construct meaning, nor does it consider the factors associated with the test (Vandenberg, 2002). In order to fully understand how theoretical models will change across groups, consideration should be made for the specifics to be measured and outside factors beyond cognitive ability that will influence test scores, none of which was done by the developers of the tests included in this study.

Given the nature of the content transition assessments seek to assess, factors not related to cognitive ability may have a powerful effect on the meaning some groups make of a construct (Cheung & Rensvold, 2002; Vandenberg, 2002; Vandenberg & Lance, 2000). For example, students with disabilities need various levels of support for participating in postsecondary education. Some students will enroll full-time and receive accommodations, while others will audit some classes part-time with support thereby changing the meaning of "postsecondary education" across groups, yet these variables typically would not be considered in the development of a measure of enrollment in postsecondary education. Instead, a student who enrolls part-time is likely to score lower, when in fact he is not doing *worse*, his meaning of postsecondary education is vastly different from that of a student who needs less support.

With these considerations in mind, simply breaking students by intelligence level does little to contribute to understanding the technical properties of transition assessments. A number of factors have been associated with successful transition from high school to postsecondary life for students with disabilities (McConnell et al., 2013; Test, Mazzotti, Mustian, Fowler, Kortering, & Kohler, 2009)—such as perseverance, and quality transition plans—and from the perspective of the transition community, it is access to these factors that are likely to contribute to inequivalence of theoretical structures across samples. This issue will be further addressed when a method for grouping participants is proposed for the *TAGG*.

Grouping Methods Not Based on Intelligence

Because the review of how students with disabilities were grouped for the development of transition assessments produced only one method, the search was expanded to examine how students with disabilities were grouped throughout special education research. The methods described are commonly found in the special education literature, particularly when instructional methods are described.

Mild-moderate/severe-profound disabilities. Generally, the category mild-moderate disabilities refers to SLD, ID with an IQ greater than 60, Speech/Language Disorders, and ED. This grouping method may drive the type of instruction a student will receive—general curriculum for students having mild-moderate disabilities, and life skills or functional skills for students with severe-profound disabilities—and placement decisions, such as whether the student is placed in the "educable mentally handicapped" class or "trainable mentally handicapped class" (Smith, Palloway, Patton & Dowdy, 1998).

Ideally, mild-moderate and severe-profound describe the degree of a student's disability in context (Luckasson et al., 1992). For example, a student who has a physical disability could be classified as having a severe-profound disability for the purposes of independent living, even though he has typical cognitive ability and average school achievement. Likewise, a student may be classified as severe-profound for having low cognitive ability, but has typical physical ability. Unfortunately, this is rarely practiced, and when students are identified as having a severe/profound disability, it is assumed the degree of their disability is global (Smith et al., 1998).

Because the mild-moderate and severe-profound labels are context dependent, using them to recruit test development participants can lead to an inappropriate sample or risk excluding students for whom the test is otherwise inappropriate. If mild-moderate and severe-profound are to be used to sort a sample into groups or for recruitment, these descriptors should be clearly defined using a preset criteria related to the context and construct being measured.

High-incidence/low-incidence disabilities. The descriptors high-incidence and low-incidence refer to the most commonly identified disabilities among children who have disabilities. Until recently, high-incidence referred to students having SLD, ED, or mild ID (Gage, Lierheimer, & Goran, 2012; Hallahan & Kauffman, 1977). Due to increases in identification, high-incidence now includes high-functioning autism, attention-deficit disorder, and speech and language disorders (Gage et al., 2012). Use of this method breaks the IDEA (2004) ID and autism categories into functional levels, and separates ADHD from other health impairment. The categories deaf-blindness, deafness, hearing impairment, multiple disabilities, orthopedic impairment, traumatic brain injury, and visual impairment are all considered low-incidence.

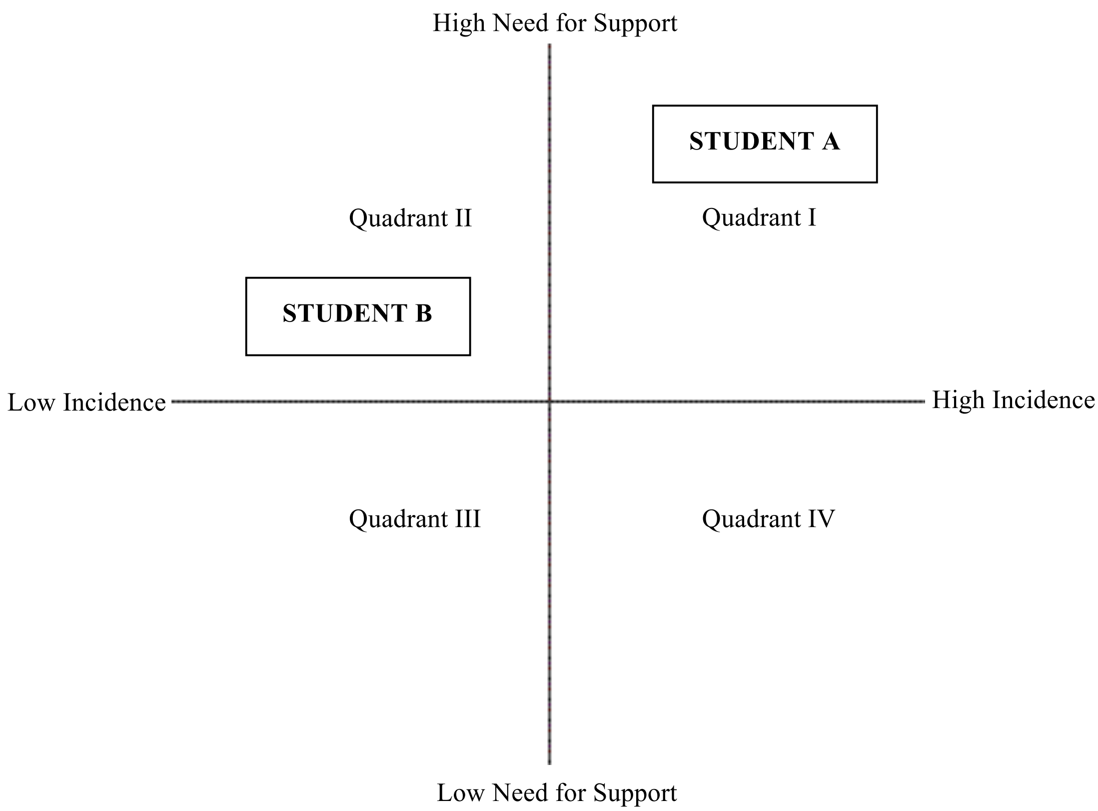
The disabilities classified as high- and low-incidence disabilities share a great deal of overlap with those classified as mild-moderate, thus when educators are told they will have a student who has a high-incidence disability, they may safely infer that the student will require moderate to little support (Gage et al., 2012). However, it cannot be assumed that because a student has a low-incidence disability he will need significant support—particularly in the cases of hearing, visual, and orthopedic impairment.

The high-incidence/low-incidence system of grouping students takes several categories of disability and combines them to make a larger category about which generalizations can be made. From a research and test development perspective, in order to capture participants who have high-functioning autism, ADHD, and mild intellectual disabilities, researchers must have access to students' medical and psychological records, which is usually not permitted. Finally, like the mild-moderate/severe-profound method of grouping students, the high-incidence/low-incidence method reinforces the notion that disabilities affect all life areas equally, and fails to consider the contexts in which a student truly has ability and a disability.

Hybrid incidence and level of need method. If ample data is collected, researchers may use a combined approach to understand and categorize their participants' disabilities. The interaction of the mild-moderate/severe-profound and high-incidence/low-incidence methods of describing students obliges the use of the twodimensional model presented in Figure 7 to develop a full picture of a student's functioning.

Figure 7

System for Grouping Students By Disability Incidence and Severity of Need



The combined method places the commonality of disabilities on the x-axis and the need for support on the y-axis. Using the coordinate plane system, Student A, who has a very high-incidence disability with severe needs falls in quadrant I, and thus, is a member of group 1. In contrast, Student B, who has a low-incidence disability and high needs falls in quadrant II, and is analyzed as a member of group 2.

This method presents challenges related to data collection and the procedures for rating students' levels of need. Specifically, data must be collected to measure participants' support needs. Because the data to be collected involves rating, procedures must be put in place to maintain high levels of inter-rater reliability (Saal, Downey, & Lahey, 1980). Furthermore, because levels of support needs vary by contexts, the rating classifications assigned using this method should be contextual, thereby limiting the generalizations that may be made about the test scores obtained using this method of classification (Messick, 1995).

Noncategorical. The noncategorical method of classifying students with disabilities has gained traction with the implementation of response to intervention identification models. This method deals principally with the classification of students whose disabilities are related directly to learning and behavior (Gage et al., 2012). Students who are found eligible for special education generally do not have disabilities to a degree that they are identified before school enrollment, thus they may be framed as mild disabilities. However, these disabilities may become more apparent, or “severe” as curriculum becomes more difficult.

The noncategorical method of grouping students with disabilities brings with it the same issues as the OHI and ED labels, it is a catch-all category that does little to describe or define its group members. Moreover, because it is a relatively new method of classifying students, and is assigned based solely on failure in academics and behavior, we have yet to understand whether there is greater meaning to be associated with this label, particularly whether there is a risk of measurement inequivalence of theoretical models for this group.

Methodological Issues Related to Grouping

Disability researchers face unique challenges with recruiting participants. Young people with disabilities are considered a special population, so researchers must be sensitive in how they approach participants and obtain assent. Because students with disabilities do not always have the communication skills to provide assent, researchers may encounter challenges obtaining permission to undertake a study from their university institutional review board.

Unfortunately, methods for examining model equivalence and test bias necessitate large sample sizes, regardless of the steps taken to complete the analysis. For Multigroup Confirmatory Factor Analysis (MGCFA), sample sizes need to range from 50 to 1,000 per group, depending on the purpose of the analysis (Marsh & Balla, 1994). Sample sizes may remain relatively low when MGCFA is used at overall model fit and item-construct levels, but samples that are too small are not sensitive enough to test for inequivalence of factor loadings (Meade & Lautenschlager, 2004). On the other hand, if MGCFA methods with sample sizes large enough to detect factor loading inequivalence are used at all levels, caution must be taken as many model fit statistics have been found to have a direct relation with sample size and model complexity, thereby underestimating model fit (Cheung & Rensvold, 2002; Marsh & Balla, 1994).

Along with the misunderstanding of the meaning of test properties (Cortina, 1993; Green, Lissitz, & Mulaik, 1977; Schmitt, 1996), these quantitative issues likely contribute to test developers' and researchers' inattentiveness in grouping students with disabilities for research purposes. Without adequate participants, test developers must resort to slipshod approaches that do not consider the meaning of the assessment, or disregard examining the assessment for model differences.

TAGG MGCFA Grouping

Because the *TAGG* is an assessment of student actions associated with postsecondary employment and education, the developers recruited participants who planned on entering postsecondary education and/or competitive employment, which were defined using Indicator 14 (NPSOC, 2013). Indicator 14 defines postsecondary education and training enrollment as part- or full-time enrollment at a community-based college or university or vocational training program with one program unit (semester, quarter, term, etc.) completed within one year of leaving high school. Competitive employment is defined as earning at or above minimum wage at a job with nondisabled peers or in self-employment for an average of 20 hours per week for a period of 90 days at any time during the year after leaving high school. Any high school student with a disability who planned on entering one or both of these types of programs was eligible for inclusion in the *TAGG* development sample, regardless of disability or degree of disability.

The *TAGG* developers collected substantial demographic data to understand the characteristics of the students, educators, and families who participated in development. These data ranged from traditional demographics such as ethnicity, disability category, and IQ, to school specific demographics they believe are associated with postsecondary employment and education, such as number of class periods in general education per school day, degree of involvement in the IEP process, and grade point average. Ultimately, the goal will be to compare the *TAGG* scores of students within these demographic groups. For example, a researcher may want to understand whether there is a relation between the degree to which students are involved in the IEP process and their *TAGG* scores. However, before the team can move forward with understanding the relation of these predictors and outcomes, they must confirm that the *TAGG* model applies to students at all levels of the predictor variables.

It is the position of this writer that the broad category Students with Disabilities overgeneralizes the findings and assumptions we make about the population. The *TAGG* was designed to be suitable for all students with disabilities who plan on entering postsecondary employment and education, so it must be reviewed for suitability across disability groups, but as this review has described, grouping students based on disability is not as simple as it may appear. Disability refers to physical and psychological characteristics of an individual, but the degree and meaning of an individual's disability is also shaped by needs, context, and opportunities (Finkelstein, 2002; Luckasson et al., 1995; Wehmeyer, 1995).

This review has found that it is impossible to understand the essence of students' disabilities using only the IDEA (2004) categorical labels or other traditional methods of grouping. All of the students included in the *TAGG* development are labeled as having a disability, but this reviewer believes that in the context of the *TAGG* framework and the traits the *TAGG* will measure, students' disabilities are a greater function of opportunities to learn and practice the *TAGG* behaviors, and not disability category. Therefore, it is proposed that the *TAGG* development team examines differences in the suitability of the *TAGG* theoretical structure not based on categorical labels, but by level of percent of time a student spends in general education.

The *TAGG* theoretical model includes eight constructs with three to six behaviors associated with each construct. Table 8 lists each of the *TAGG* constructs, along with examples of the measured behaviors associated with each construct.

Table 8

TAGG Constructs and Examples of Associated Behaviors

Construct	Associated Behaviors
Strengths and Limitations	Tells what he does well Identifies when support is needed
Disability Awareness	Uses least stigmatizing disability label* Explains to other why he receives special education services*
Persistence	Values not giving up Learns from mistakes and improves in next attempt
Interacting with Others	Sets goals based in interests and input from others* Has attained at least one IEP transition plan goal*
Goal Setting and Attainment	Breaks goals into smaller parts* Completed at least one IEP annual transition goal
Employment	Expresses wanting a job Has had a paid job
Involvement in the IEP	Told IEP team goals* Led IEP meeting*
Support Community	Distinguishes between positive and nonpositive support people Seeks assistance from community agencies

Note. * indicates special education curriculum developed for instruction.

Some of the *TAGG* constructs specifically address disability-specific issues I believe students are unlikely to learn in the general education setting (i.e., disability awareness, and IEP involvement), while others include skills students may improve through practice in the general education setting (i.e., utilization of supports, and interacting with others). The range of settings in which the *TAGG* behaviors may be learned and practiced, and the variance within the *TAGG* sample in the percent of time spent in special education and general education settings provide a second rationale for examining the suitability of the *TAGG* for use with all students with disabilities. (The first rationale being that type of disability is not expected to act as a significant unintended factor on *TAGG* scores.)

Finally, sound methodology for any multigroup analysis of bias or model fit necessitates a minimum of about 50 participants per group (Meade, Johnson, & Braddy, 2006). To meet that condition, the *TAGG* would need 50 students in each disability group who planned on entering postsecondary education and competitive employment, which have not been collected. The review was completed, in part, to develop a method for grouping students with disabilities, and found that even clustering disability categories presents significant problems, largely due to poorly defined inclusion or exclusion criteria for the 13 IDEA (2004) disability categories.

This review presented various methods for grouping students with disabilities. Based on this review and the preceding sections of this manuscript, a strong need has been established to examine the equivalence of the *TAGG* structure across students with various disabilities. This review takes the position that the best method for grouping *TAGG* participants by disability is to infer that students who have greater needs spend greater time in special education, while those who have minimal needs spend more time in general education (Griffin, Taylor, Urbano, & Hodapp, 2013). Although Griffin et al. focused on students with autism, it is this author's perspective that the same is true for all students with disabilities.

Because opportunities to learn and act on the *TAGG* behaviors vary across the general education setting, and degree of disability may be related to participation in general education, this researcher believes that the *TAGG* factor structures for skills taught in special education will not be equivalent for students who do not regularly attend special education. Likewise, it is expected that the *TAGG* factors and behaviors that are likely to be developed in general education will not hold among students who did not participate in general education.

Up to this point, this manuscript has presented an argument for the use of multigroup factor analysis procedures to understand the suitability of a test's theoretical structure across students with disabilities. Historically, these procedures have not been used, and it has been assumed that if a test is designed for students with disabilities, it is thus suitable for all students with disabilities. Furthermore, it is unlikely that simply having a disability makes a student different enough from a non-disabled peer that a test needs to be designed exclusively to understand him. That said, educators do, in fact, use assessments designed for students with disabilities, and those assessments must be evaluated for their quality.

Understanding the quality of an assessment involves examining the relation and interaction between the various facets of validity and reliability. Although disability and accommodations contribute to test scores, other factors, such as knowledge of the constructs being assessed, likely contribute to test scores more than the condition of having a disability. Therefore, this study will examine how participation in general education influences the meaning students, families, and educators make of the *TAGG* constructs.

CHAPTER THREE

Methodology

This study sought to understand whether the structure of the *TAGG* changes as students spend more or less time in general education due to having a disability. Presently, the *TAGG* scoring profile generates goals and assumes that all goals are of equal value to all students with disabilities, regardless the extent of their participation in special and general education.

This study sought to test this assumption by first assessing whether the items were associated with the same constructs at various degrees of general education participation. For example, this study wanted to assure that if the *TAGG* suggested a goal that a student would make a plan for attaining his transition goals, that goal would remain in the Goal Attainment construct and would not shift construct groups depending on level of participation in general education.

To do this, the study examined the fit of the *TAGG* theoretical model across the sample of *TAGG* students grouped by percent of time spent in general education. The study also examined whether the *TAGG* theoretical structure was suitable across grade levels and exposure to transition education. Finally, the study, out of special education tradition, examined the fit of the *TAGG* structure across disability categories.

Research Questions

This study sought to answer the following questions for various groups of students, including students who spend more or less time in general education, students in different grades, students who have had some transition education, and students in some disability categories:

1. To determine if configural invariance exists, are the items on the *TAGG* associated with the same constructs regardless of the percent of time students spend in general education, amount transition education, grade level, or disability group?
2. To determine if metric invariance exists, do the items on the *TAGG* have the same strength of relations to their constructs regardless of the percent of time students spend in general education, amount transition education, grade level, or disability group?

Data Collection

The dataset for this study was collected in three annual phases as part of the development of the *TAGG*. Each *TAGG* case represents one student participant. Each case includes, at a minimum, a *TAGG* professional version, a *TAGG* student version, and professional and student demographics completed by the participating professional. Many cases also include *TAGG* family versions and family demographics, and self-evaluation checklists for fidelity of administration purposes.

Recruitment. Recruitment for participation in the *TAGG* began by contacting transition professionals using a transition focused email list developed from state transition institutes and national conferences. Transition educators and professionals who expressed interest were sent a link to a training video. After watching the training video, educators could agree to participate.

Transition professionals who agreed to participate were asked to refer high school students who they believed had mild/moderate disabilities during the first two phases of *TAGG* development. During the third phase development, the phrase "...and who plan on entering postsecondary education and/or competitive employment" was added to clarify the intended population.

Settings. *TAGG* development data were collected from educators, students and families at 120 schools in 30 states from 2010 through 2013. Educators worked mostly in public schools (95%), and a few worked in private schools (< 1%) and charter schools (4%). Table 9 presents the number and percent of professionals, families, and students who returned *TAGG* forms from each state.

Table 9

Percent of TAGG Participants by State

	Professional[a]		Family		Student	
State	n	%	n	%	n	%
Alabama	17	1%	12	2	17	1
Alaska	10	< 1	1	< 1	10	< 1
Arizona	49	3	17	2	46	3
Arkansas	222	15	113	16	219	15
California	36	2	17	2	36	2
Colorado	191	13	139	20	194	13
Delaware	21	1	3	< 1	20	1
Idaho	2	< 1	0	0	2	< 1
Illinois	64	4	31	4	43	4
Indiana	9	< 1	8	1	6	< 1
Iowa	12	< 1	8	1	12	< 1
Kansas	1	< 1	1	< 1	0	0
Missouri	24	2	16	2	25	2
Nebraska	5	< 1	0	0	5	< 1
Nevada	6	< 1	1	< 1	6	< 1
New Hampshire	2	< 1	0	0	2	< 1
New Jersey	10	< 1	0	0	10	< 1
New Mexico	82	5	45	6	83	6
New York	10	< 1	3	< 1	10	< 1
North Carolina	63	4	36	5	60	4
Ohio	22	1	4	< 1	21	1
Oklahoma	462	31	385	26	458	31
Pennsylvania	4	< 1	3	< 1	4	< 1
Rhode Island	21	1	15	2	21	1
South Carolina	13	< 1	8	1	13	< 1

Table 9

Percent of TAGG Participants by State, continued

	Professional		Family		Student	
State	n	%	n	%	n	%
Tennessee	11	1	2	< 1	10	< 1
Texas	18	< 1	0	0	18	1
Virginia	3	< 1	2	< 1	3	< 1
West Virginia	55	4	30	4	53	4
Wisconsin	69	5	13	2	64	4
Total	1514		913		1471	

Note. [a]Professionals returned forms for multiple students.

***TAGG* administration procedures.** Prior to administering the *TAGG*, professionals watched a 15-minute training video that described the purpose of the project and their responsibilities as participants. The professionals also received a consent form that described their duties, an instruction sheet, and a self-evaluation checklist.

Before administering the *TAGG*, professionals obtained a signed agreement from their building principal or program director giving permission for professional educators and students to participate. Once permission was obtained from the principal, students' families were contacted to obtain consent for the student to participate and to recruit families for participation. Next, demographic forms were completed for each student and professional along with *TAGG* professional versions for each student. Finally, the student version of the *TAGG* was administered, and the professional coordinated administration of the *TAGG* family version and family demographics. Materials were then returned to the research center.

Collection of demographic data. Professionals completed demographic forms for themselves and for each student who participated in the development process.

Family members who participated in *TAGG* development were also asked to complete demographic forms. The professional demographic forms asked for traditional personal characteristic information such as age, number of years of teaching experience, and ethnicity, along with questions about training in transition education and assessment, and use of transition assessments and instructional material. Student demographic sheets asked educators to indicate the student's age, gender, ethnicity, disability and grade, along with placement information and his participation in the IEP process. Family demographic forms asked about ethnicity, language spoken in the home, work status, and whether help had been obtained for completing the *TAGG* family version assessment.

***TAGG* student version administration.** Professional participants administered the *TAGG* individually or in small groups using either an online website or paper versions. Educators read the directions and purpose of the assessment to the students, and asked the students if they had any questions. The students were told they could ask questions and were reminded to think before responding to items. Finally, the educators were asked to check the assessments for missed items when they were collected. If a student missed an item, the educator was to ask the student if he wanted to respond to it.

***TAGG* family version administration.** Each *TAGG* family version included a cover letter explaining the purpose of the assessment and directions for returning the assessment to their student's teacher. Family members were asked to think about how often, on a scale of 1-5, their student had exhibited the *TAGG* behaviors. Spanish language forms were available for families, who were encouraged to ask their student's teacher for clarification if they did not understand.

Honorarium. Participants could complete *TAGG* assessments and demographics online or with paper versions. Each professional who participated received a $30 honorarium for each completed case they returned. Family members and students each received a $10 gift card when they completed *TAGG* assessments.

Fidelity of Administration and Data Entry

Professional educators were asked to complete an administration checklist for each *TAGG* administration. The checklist included seven yes or no items indicating whether they had implemented each step in administration. For the first two phases of data collection, educators implemented 98.8% of the instructions, with a range of 86% (yes to six of the seven items) to 100%. For group administrations, educator indicated following *TAGG* administration procedures 97.7% of the time, with a range of 86-100%, and for individual administrations, educators indicated following procedures with 100% accuracy. No data has been analyzed on the on fidelity of administration for the third phase of data collection.

Field observations. Field observations of 10 professional educators were completed during the first phase of data collection. The field observations used the same checklist the professional educators used, and then was compared with the checklist collected from the professional who was observed administering the *TAGG*. Professional educators who were observed reported following the administration procedures with 98.8% accuracy. When compared to the checklists returned by the field observer, there was 97% agreement between the administrating professional and field observer.

Field observations of 10 educators were also completed during the second phase of data collection. The professional educators indicated they implemented 92.3% (82-100%) of the administration procedures. The field observer indicated the professional educators followed the procedures with 90% (71-100%) accuracy, indicating one disagreement.

Data entry. Multiple research assistants entered and checked a portion of the data set. Agreement for the first phase of data collection was 99.6%, for the second phase agreement was 99.8%, and for the third phase agreement was 99.5%. Disagreements were addressed using a consensus discussion at research team meetings.

Participants

Data was collected from 30 states, with 139 professional educators from 120 schools completing *TAGG* professional versions for 1,537 students who were ages 14-21 and in grades 9-12 or programs for students ages 18-21. Students completed 1,418 student versions, and 793 families returned family versions. Educators indicated the number of periods in each student's school day and the number of periods the students spent in general education for a total for 1,363 students. Seven of those students were removed because the percent of their day spent in general education was calculated to be greater than 100%, leaving a total of 1,452 completed professional *TAGG* versions, 1,356 completed student *TAGG* versions, and 793 family *TAGG* versions. The demographic information below describes the professionals, students, and family members who participated in *TAGG* development.

Students from each of the disability categories were represented, with 17 students having no disability category specified. The students had spent an average of

68% of their school day in the general education setting. Fifty-six students received English second language (ESL) services. Of the parents who completed *TAGG* family versions, 23 indicated they did not work due to a disability, although the actual number is likely greater as this information was only collected for phase three.

Sampling issues. Some important differences in the three samples were considered in the development of this study. First, the criteria for inclusion in *TAGG* development evolved across the three phases of data collection, particularly due to the addition of the expression "…and who plan on entering postsecondary education and/or competitive employment." This statement may skew the sample toward a lower degree of overall need, or exclude students who fit the intended criteria for inclusion, but who do not express plans for postsecondary education and/or competitive employment. This may have also resulted in educators under-referring students who they believed had the capacity for entering postsecondary education and competitive employment, but who lacked motivation.

Second, the second phase of data collection produced a sample of students who were, on average, about a year younger than the first phase's sample. This has implications for the construct definitions and the meaningfulness of the construct definitions to the test takers. A factor analysis indicated the Employment construct was not as meaningful to this group as the prior group, likely because some of these students were not yet eligible for paid employment.

Finally, the first phase of data collection produced a group of students who, based on their academic enrollment, appeared to have disabilities that were more severe than intended for data collection. One group of students referred by one educator was removed from the overall data set because they appeared to be enrolled in classes geared exclusively toward life skills education, but there is no way to ensure that other similar instances were not overlooked. Additionally, the first year of data collection used the development version of the *TAGG*, which included 75 items rather than just the final 34 items.

Participant demographics. This study used data from the three annual phases of *TAGG* development. The decision about which phases to include was made based on the acceptable fit of the *TAGG* model across all three phases of data collection and all three versions of the *TAGG*. Tables 10-12 present the demographics for professionals, families, and students who participated in this study. On average, professionals completed about 10 *TAGG* professional versions for 10 of their students.

Table 10

Full TAGG Sample Educator Participants (n = 139)

Characteristic		n	%	Missing[a]
Age				15%
18-24		5	4	
25-29		7	6	
30-34		6	5	
35-39		12	10	
40-44		17	14	
45-49		14	12	
50-54		22	19	
55-59		22	19	
> 60		13	11	
	Mean	45.4		
	SD	10.4		
Gender				
Male		8	6	0%
Female		131	94	

Table 10

Full TAGG Sample Educator Participants (n = 139), continued

Characteristic		n	%	Missing[a]
Level of Education				1%
Bachelor's Degree		18	13	
Some Master's Coursework		31	23	
Master's Degree		72	53	
Educational Specialist		8	6	
Some Doctoral Coursework		7	5	
Ph.D. or Ed.D.		1	< 1	
Years teaching students with disabilities				0%
< 3		10	7	
3-10		52	37	
> 10		77	55	
	Mean	15.0		
	SD	10.3		
Years teaching at current school				0%
< 3		37	27	
3-10		62	45	
> 10		40	29	
	Mean	8.3		
	SD	7.0		
Years teaching in current district				0%
< 3		68	49	
3-10		41	30	
> 10		30	22	
	Mean	10.9		
	SD	9.0		
Lives in district		70	51	1%

Note. [a]Missing includes "Not indicated" and "Don't Know".

Table 10

Full TAGG Sample Educator Participants (n = 139), continued

Characteristic	n	%	Missing[a]
Number of students on caseload			
< 5	10	7	
5-10	10	7	
10-15	29	21	
> 15	90	65	
Mean	31.2		
SD	49.2		
Number of professional development days in transition education			0%
< 10	97	70	
10-20	21	15	
> 20	21	15	
Ethnicity			0%
White or Caucasian	115	83	
Black or African-American	12	9	
American Indian or Alaska Native	6	4	
Mexican, Mexican-American, or Chicano	1	< 1	
Hispanic or Latino	3	2	
Asian	2	1	

Note. [a]Missing includes “Not indicated” and “Don’t Know”.

Table 11

Full TAGG Sample Family Participants (n = 793)

Characteristic	n	%	Missing[a]
Age			0%
< 18	68	9%	
18-24	16	2	
25-29	8	1	
30-34	26	3	
35-39	152	19	
40-44	200	25	
45-49	148	19	
50-54	81	10	
55-59	49	6	
> 60	45	6	
Mean	44.3		
SD	9.2		
Primary Language			< 1%
English	728	93%	
Other	56	7	
Relationship to Student			1%
Father	98	13%	
Grandfather	7	< 1	
Brother	2	< 1	
Male Guardian	4	< 1	
Other	10	1	
Mother	598	76	
Grandmother	32	4	
Aunt	4	< 1	
Sister	4	< 1	
Female Guardian	14	2	
Stepfather	3	< 1	
Stepmother	6	< 1	

Note. [a]Missing includes “Not indicated” and “Don’t Know”.

Table 11

Full TAGG Sample Family Participants (n = 793), continued

Characteristic	n	%	Missing[a]
Education			3%
Less than high school	104	14%	
High school diploma/GED	310	40	
Vocational certificate	92	12	
Associate's degree	111	14	
Bachelor's degree	111	14	
Master's degree	30	4	
Doctorate/Professional degree	9	1	
Employment			< 1%
Employed full-time	405	51%	
Employed part-time	70	9	
Self-employed full-time	35	4	
Self-employed part-time	20	3	
Not working	187	24	
Retired	48	6	
Permanently disabled, not working for pay	23	3	
Received help completing *TAGG*	56	7%	
Help received			
Reading	12	35%	
Writing	3	9	
Translating	4	12	
Explanation	12	35	
Other	3	9	

Note. [a]Missing includes "Not indicated" and "Don't Know".

Table 11

Full TAGG Sample Family Participants (n = 793), continued

Characteristic	n	%	Missing[a]
Ethnicity			
White	442	65	
Black or African-American	66	10	
American Indian or Alaska Native	664	9	
Mexican, Mexican-American, Chicano	34	5	
Puerto Rican	5	< 1	
Cuban	2	< 1	
Hispanic or Latino	49	7	
Asian	4	< 1	
Other (Unidentified)	1	< 1	
Egyptian	2	< 1	
Spanish	1	< 1	
German	1	< 1	
Multiethnic	3	< 1	

Note. [a]Missing includes “Not indicated” and “Don’t Know”.

Table 12

Full TAGG Sample Student Participants (n = 1,356)

Characteristic		n	%	Missing[a]
Age				8%
14		31	2	
15		94	7	
16		172	12	
17		352	25	
18-21		749	54	
	Mean	17.4		
	SD	1.29		
Gender				< 1%
Male		853	59	
Female		592	41	
Primary Disability				1%
Autism		31	4	
Deaf-Blindness		2	< 1	
Emotional Disturbance		79	6	
Hearing Impaired		15	1	
Intellectual Disability		174	12	
Multiple Disability		16	1	
Orthopedic Impairment		11	< 1	
Other Health Impairment		200	14	
Specific Learning Disability		828	58	
Speech/Language Disability		21	1	
Traumatic Brain Injury		9	< 1	
Visual Impairment/Blindness		5	< 1	
Other		14	1	
Level of Disability				8%
Mild-Moderate		1280	90	
Severe-Profound		64	4	
Attendance				2%
Absent ≤ 1 day/month		848	60	
Absent 2-3 days/month		410	29	
Absent ≥ 4 days/month		164	12	

Note. [a]Missing includes "Not indicated" and "Don't Know".

Table 12

Full TAGG Sample Student Participants (n = 1,356), continued

Characteristic		N	%	Missing[a]
Attended IEP meeting		1267	88	15%
Actively participated in IEP meeting		699	62	32%
Led IEP meeting		143	8	9%
Years teacher has known student				48%
< 1 year		274	34	
1-2 years		233	29	
2-3 years		182	23	
3-4 years		34	4	
4-5 years		49	6	
> 5 years		27	3	
	Mean	2.6		
	SD	2.3		
Years student has attended school				29%
< 1 year		198	19	
1-2 years		207	20	
2-3 years		227	21	
3-4 years		254	24	
4-5 years		171	16	
	Mean	3.2		
	SD	2.5		
Years student has been enrolled in district				18%
< 1 year		142	12	
1-2 years		66	5	
2-3 years		68	6	
3-4 years		28	2	
4-5 years		302	25	
> 5 years		134	14	
	Mean	7.9		
	SD	4.5		
Eligible for free/reduced lunch		824	58	18%
Receives free/reduced lunch		787	58	8%
Receives ESL services		56	4	

Note. [a]Missing includes “Not indicated” and “Don’t Know”.

Table 12

Full TAGG Sample Student Participants (n = 1,356), continued

Characteristic	N	%	Missing[a]
Ethnicity			16%
American	1	< 1	
American Indian or Alaska Native	121	10	
Arabic	1	< 1	
Asian	9	< 1	
Black or African-American	245	20	
Cuban	2	< 1	
Egyptian	1	< 1	
Guamanian, Chamorro	1	< 1	
Hispanic or Latino	105	9	
Irish	1	< 1	
Multiethnic	3	< 1	
Mexican, Mexican-American, or Chicano	81	7	
Native Hawaiian	1	< 1	
Other (Not Indicated)	6	< 1	
Puerto Rican	5	< 1	
White or Caucasian	646	53	

Note. [a]Missing includes “Not indicated” and “Don’t Know”.

Instrumentation

All three versions of the *TAGG* are divided into eight constructs, which were developed through a review of the transition literature. The construct definitions describe the behaviors associated with each construct. It is from these behaviors that goals will be generated when each score profile is produced. Table 8 presents examples of the behaviors included on each version of the *TAGG*, and summaries of the construct definitions are described below.

Knowledge of strengths and limitations. The student expresses areas in which he is strong or limited. He knows how his strengths and limitations affect different situations.

Disability awareness. The student knows he has a disability and can describe his needs. He can explain how his disability affects his life, and supports he needs and is legally allowed.

Persistence. The student believes he can surpass challenges by spending extra time or effort. He also sees failures as chances to learn.

Interacting with others. The student participates with other individuals in community and school settings.

Goal setting and attainment. The student breaks long term goals into short-term goals. He makes a plan to attain his goals. When the student does not attain his goal, he revises and adjusts his plan.

Employment. The student has had a paid job, and he has expressed wanting a job that matches his interests and abilities.

Involvement in the IEP. The student actively participates in his annual Individual Education Plan meetings. He discusses his current performance and plans to attain his academic goals.

Community Supports. The student uses resources and positive support people when appropriate.

***TAGG* Technical Properties**

This study used all three versions of the *TAGG* (Martin et al., 2011)—Professional, Family, and Student, which were each subjected to factor analysis, tests of internal consistency using Cronbach's α (Cronbach, 1951) for each of the constructs, and test-retest reliability.

All three versions of the *TAGG* had moderately to highly correlated constructs, indicating that high ability in one construct is related to high ability in the other constructs (Hennessey et al., 2013). The information listed below reflects the analysis of the first and second phases of *TAGG* development, as the third phase of development has not yet been included in the *TAGG* technical manual.

***TAGG* professional version.** The *TAGG* professional version was developed to be completed by any transition professional including special educators, rehabilitation counselors, and job coaches who spend time working directly with the student. The *TAGG* eight-construct structure had acceptable fit for the *TAGG-P* (χ^2 = 2863.49, *df* = 1021, RMSEA = .072, CFI = .88, TLI = .88, RMSR = .065) for phase two of data collection. Internal consistency for the *TAGG* professional version ranged from .68-.93 for the individual constructs for the first phase of data collection and .64-.94 for the second phase of data collection. Finally, a test-retest administration indicated a large correlation between administrations (.80) of the *TAGG* professional version, following a three-month delay (Hennessey et al., 2013).

***TAGG* family version.** The *TAGG* family version was developed to be completed by any family member who spends time with the student. The *TAGG* eight-construct structure also has acceptable fit for the *TAGG-F* (χ^2 = 1995.76, *df* = 1087, RMSEA = .0579, CFI = .89, TLI = .89, RMSR = .0679) for phase two of data collection. Internal consistency for the *TAGG* family version ranged from .52-.91 for the individual constructs for the first phase of data collection and .52-.90 for the second phase of data collection. Finally, a test-retest administration indicated a moderate correlation between administrations (.70) of the *TAGG* family version (Hennessey et al., 2013).

***TAGG* student version.** The *TAGG* student version was developed to be completed by the student. The *TAGG* student version seven-construct structure has excellent fit (χ^2 = 1879.42, *df* = 1028, RMSEA = .0490, CFI = .87, TLI = .86, RMSR = .0762). Internal consistency for the *TAGG* student version ranged from .44-.82 for the individual constructs for the first phase of data collection and .45-.82 for the second phase of data collection. Finally, a test-retest administration indicated a moderate correlation between administrations (.70) of the *TAGG* student version (Hennessey et al., 2013).

Data Analysis Procedures

***TAGG* score computation.** Scores for the individual profiles generated by the *TAGG* were computed using item response theory methods. However, in their review of the MGCFA literature, Vandenberg and Lance (2000) put forth a list of assumptions to be met before multigroup invariance can be examined. With regards to scoring, Vandenberg and Lance recommend construct scores be developed using additive methods, therefore, *TAGG* construct scores for this study will be developed by adding the scores for each item within a construct. For example, if a student is rated 3, 4, and 5 for a set of three items on a construct, his score will be calculated by adding 3 + 4 +5 for a total score of 12 on the construct. To replicate the CFA procedures used for phases one and two of the *TAGG*, binary yes/no items will be scored a zero for no, and five for yes on the educator and family forms and one for yes on the student form.

Missing data TAGG items. This study faced two issues related to missing data. First, the research plan for the *TAGG* called for the completion of one family *TAGG*, one educator *TAGG*, and one student *TAGG* for each student participant. Family participation was optional, resulting in incomplete data for some students. This reduced the overall sample size, and thus the sizes of the subgroups when the family *TAGG* version was analyzed.

The second issue related to missing data involved items participants chose not to answer when completing the *TAGG*. Scores for Likert-type items were imputed using each student's mean item score for the construct rounded to the nearest whole number. For example, if the student had scores of 4, 4, 5, and 4 plus one missing item on a construct having five items, the missing item was calculated as (4 + 4 + 5 + 4) / 4 = 4.25. The score was then rounded to 4 to be consistent with other *TAGG* item scores and substituted for the students missing score. Finally, the student's construct score was calculated by adding all of the scores, plus the imputed score.

This method had some limitations, however. The *TAGG* constructs range in number of items. One construct has three items (Interacting with Others), most have four or five items (Strengths and Limitations, Disability Awareness, Persistence, Employment, Community Supports, Employment, and Involvement in the IEP), and one construct has six items (Goal Attainment). For the constructs having only three or four items, no more than one item score was imputed. When the constructs had five or six items, not more than two item scores were imputed (Shrive, Stuart, Quan, & Ghali, 2006). Furthermore, because binary items are not scored on the same Likert-type scale as the other *TAGG* items, they could not be imputed.

This method presented challenges for the Employment construct, which has four items, two of which are binary, so no imputations could be made for this construct. One binary item in the Goal Attainment construct on the *TAGG-F* was also problematic as nearly half of the families who participated did not respond to the item and the item could not be imputed. Fortunately, the Goal Attainment construct includes five other items.

Calculation of percent of time in general education. The *TAGG* demographic data forms included two items that will be used for the computation of the percent of time students spent in general education. The first demographic item asked teachers how many class periods there were in the student's school day (number of class periods per day). The second item asked educators the number of class periods the student spent in general education (number of periods in general education).

The percent of time spent in general education was calculated by dividing the number of periods in special education by the number of class periods per day and then multiplying by 100. For example, if a student has six total class periods, and he spent four of those periods in general education, his percent of time spent in general education was calculated as (4 / 6)100 = 67%.

Missing demographic data. The primary demographic data necessary for completing this study are the number of periods in the students' school day and the number of periods he spends in special education. Without these, no grouping variable could be completed, and the case was dropped from the full data set. For example, a total of 1,418 *TAGG-S* forms were completed, but the percent of school day in general education could only be calculated for 1,356 students so 62 cases were dropped from analysis. Similarly, when groups were developed for the transition education groups, when professional educators indicated they did not know if a student had instruction in the skills, no level of transition education could be calculated and the case was removed. When no grouping variable needed to be calculated, such as for grade or disability category, the case was retained and the student was left as a part of the non-focus group.

Research Design

This study used a multigroup confirmatory factor analysis (Cheung & Rensvold, 2002) to understand whether there were differences in the *TAGG* structure across students who spent more or less time in general education, had some or no transition education, were in different grades, or were in different disability groups. The MGCFA process used a nested model described in Table 13 to compare changes in model fit as restrictions were placed on the various portions of the model (Cheung & Rensvold, 2002; Milfont & Fischer, 2010).

Table 13

Measurement Invariance Tests

Research Question	Purpose	Description	Significance Tests: Fit	Significance Tests: Change in Model Fit
1	Same items are associated with same latent traits for each group.	Version (P, F, or S) structure applied to each group	RMSEA (< .08), SRMR (< .08), CFI (> .95)	Not applicable
2	Items have same factor loadings for each group.	Factor loadings for each group are constrained to be equal to those of the full sample.	RMSEA (< .08), χ^2 (likelihood ratio), SRMR (< .08), CFI (>.95)	ΔCFI ΔMc ΔGamma^

Note: ΔCFI , ΔMc, and ΔGamma^ are described in the Data Analysis section.

Research question one. *TAGG* development has been shaped by the eight–factor structure that emerged from the data collected on the professional version for item analysis in Phase I. That structure was tested for acceptable fit on the family and student versions. The fit of the eight-factor structure from the professional form was acceptable for the family form, but not for the student form, which produced a seven-factor structure. Although some of the items on the family and student versions did not demonstrate a statistically meaningful relation to their constructs, they were retained for uniformity across the three *TAGG* versions.

Substantially more data has been collected since the two factor structures were developed, and the purpose for examining the factor structure has evolved. Initially, the purpose was to develop an instrument that would generate scores that were comparable, which necessitated a common structure. Now, the purpose for examining the structure is to understand which behaviors may be recommended as goals for students to work on for the next year they are in school.

Doing so indicates revisiting each *TAGG* version to ensure that the behaviors the *TAGG* scoring profile recommends are indeed behaviors endorsed by the test users, particularly given "less than desirable" fit of the professional *TAGG* structure to the family data when the strengths of the relations of the item to the constructs was set to be equal to those of the professional educators (χ^2 = 2681.6678, *df* = 1121, RMSEA = .075, CFI = .82, TLI = .82, RMSR = .151; Hennessey et al., 2013). For example, the present *TAGG* structure indicates that for all three versions of the *TAGG*, a suggested goal in the Interacting with Others construct might be for the student to improve his relations with teachers, family, and other adults. However, when the construct-behavior associations are examined for the family and student versions of the *TAGG*, that behavior may be associated with some other construct. Should this happen, the algorithm for suggesting goals needs to be revised to reflect the construct under which a goal is suggested.

Examining the factor structure that emerges for each version of the *TAGG* also circumvents a greater threat to the valid use of test scores. The *TAGG* claims to suggest goals based on a theoretical model of beliefs about the actions students should take in school if they are going to enter postsecondary education and/or competitive employment. Including the family and student versions of the *TAGG* implies that the belief values of those groups have been assessed and will be applied in the development of the goals. Unfortunately, only the professional *TAGG* data was used to build the list of *TAGG* items from which the goals will be drawn, so it must be understood the *TAGG* model is in fact a model of *professional educator* beliefs and the acceptability of those beliefs for families and students. In practice, this may translate to conflict within the IEP team if a goal is suggested that does not represent what families and students believe are necessary skills. In order to promote the *TAGG* as an assessment that uses data from multiple perspectives, the item-construct relations must be examined for each version to ensure that appropriate goals are generated.

Therefore, the first research question was used to assess whether the item-construct associations that emerge from each version of the *TAGG* are suitable across groups of students at various degrees of participation in general education (Cheung & Rensvold, 2002). In technical terms, this step forced the items to load the same way across the full version sample and group of examination, but allows the factor loadings to be unequal and then examines the model fit. Stated plainly, this step sought to understand whether the same items were associated with the same constructs in both the group of focus (for example, 9^{th} grade students who took the *TAGG-S*, the professional educators who completed the *TAGG-P* on 9^{th} grade students, and family members who completed the *TAGG-F* on their 9^{th} grade students) and the remaining *TAGG* participants (e.g. students in grades 10-12 and 18-21 year old programs). In practical terms, this step exposes whether the goals generated by the *TAGG* in each of the construct areas need to be reassigned to different construct areas depending on a student's group membership.

Research question two. The second research question was used to understand whether each version of the *TAGG's* items had the same strength of relation to the constructs for each focus group as for the other *TAGG* participants. For example, students who spend a great deal of time in special education may view leading their IEP meeting as more valuable to having a high level of Involvement in the IEP than students who spend little time in special education. In practical terms, this step helped understand whether any of the groups had a unique set of views about the degree to which the *TAGG* behaviors are related to or indicative of the *TAGG* constructs. To do this, factor loadings for the each subgroup were set to be equal to those estimated for the full *TAGG* sample (Cheung & Rensvold, 2002).

This step of the analysis used change in model fit statistics to understand whether applying the strength of item-construct relations improved or decayed the fit estimated to answer the first research question, but rather than using the RMSEA, SRMR, and CFI, this step used the change in CFI (ΔCFI), change in McDonald's NFI (ΔMc), and change in Gamma^ (ΔGamma^) to assess significance. If these change statistics indicated a significant decay in model fit, then the process for prioritizing which goals are suggested on the scoring profile needs to be modified to reflect how the grouping variable (time in general education, grade level, etc.) influenced which goals are of greatest priority. If the change statistics indicate the model did not deteriorate significantly by making the strength of the item-factor relations for the subgroup equal to those of the full *TAGG* sample, then the degree to which a student participates in general education does not need to be considered when goals are suggested.

Data Analysis. The multigroup analysis used for this study considered all three versions of the *TAGG*, with the fit of the structure of each version being considered for disability category, grade level, participation in general education, and exposure to transition education concepts. For the three *TAGG* versions, each group was analyzed independently using the two step process described above, for a total of two analyses per group per form, producing a total of 408 sets of fit statistics and 153 significance tests. Table 14 lists the groups to be analyzed and a description of each group.

Table 14

Number of Participants Per Group Across TAGG Versions

Group	*TAGG-P* n	*TAGG-F* n	*TAGG-S* n	Description
< -1.5 SD	139	63	123	Students who spend less than 24.9% of their day in general education
-1.5 SD	114	58	103	Students who spend 25% to 39.2% of their day in general education
-1 SD	140	76	131	Students who spend 39.3 to 53.6% of their day in general education
-.5 SD	165	76	157	Students who spend 53.7 to 67.9% of their day in general education
.5 SD	302	168	281	Students who spend 68.0 to 82.3% of their day in general education
1 SD	268	157	255	Students who spend 82.4 to 96.6% of their day in general education
> 1 SD	320	159	302	Students who spend 96.7-100% of their day in general education
Some transition education	405	252	395	Student who have had disability awareness instruction, and instruction in leading their IEP meetings or actively participating in their IEP meetings

Table 14

Number of Participants Per Group Across TAGG Versions, continued

Group	*TAGG-P* n	*TAGG-F* n	*TAGG-S* n	Description
Little to no transition education	333	141	315	Student who have had disability awareness instruction, or instruction in leading their IEP meetings or actively participating in their IEP meetings, or no transition education.
9th Grade	124	105	117	Students who took the *TAGG* in the 9th grade.
10th Grade	162	122	152	Students who took the *TAGG* in the 10th grade.
11th Grade	480	289	429	Students who took the *TAGG* in the 11th grade.
12th Grade	628	241	587	Students who took the *TAGG* in the 12th grade.
18-21 Program	74	34*	67	Students who took the *TAGG* while attending a program for students ages 18-21.
Autism	61	45*	55	Students whose educators indicated the students have autism
ID	174	97	154	Students whose educators indicated the students have intellectual disabilities
OHI	200	109	189	Students whose educators indicated the students have other health impairments

Table 14

Number of Participants Per Group Across TAGG Versions, continued

Group	*TAGG-P* n	*TAGG-F* n	*TAGG-S* n	Description
ED	79	42*	71	Students whose educators indicated the students have emotional disabilities
SLD	828	442	778	Students whose educators indicated the students have specific learning disabilities

Note. *No analysis will be completed on groups having fewer than 50 participants as the sample size may not be adequate for detecting differences in the item-construct relations (Marsh & Balla, 1994).

Significance Tests. MGCFA designs use a multi-step process which tests for improvement of the model fit as the structure of an individual group is forced to be progressively more like the full sample. The first step of the analysis assessed the fit of the *TAGG* structure on each of the subgroups when same item-construct relations were set for both the subgroup and the *TAGG* sample. Acceptable fit for this step was assessed using the RMSEA, SRMR, NNFI, and CFI (Hu & Bentler, 1999).

Because MGCFA uses nested models, the statistics used should not be correlated with the overall fit of the model or with the statistics generated for prior steps in the analysis (Cheung & Rensvold, 2002). Many fit indices have been developed for a variety of purposes, but this study will use three alternative fit indices—the change in Comparative Fit Index (ΔCFI), the change in McDonald's Non-centrality Index (ΔMc), and the change in Gamma-hat (ΔGamma^). These indices are independent of the overall fit of the *TAGG* structure to the model, thereby providing two ways to test the suitability of the model—the first way examining for improvement or deterioration of the model fit, and the second way examining the overall model fit (Cheung & Rensvold, 2002; Meade, Johnson, & Braddy, 2006; Pirritano, 2005). Finally, Meade, Johnson, and Braddy point out that when small sample sizes exist, as is the case with this study, use of alternative fit indices, rather than likelihood ratio tests (e.g. chi-square tests), reduces the influence of sample size. Use of alternative fit indices also deals with the unequal group sizes experienced in this study.

The process for calculating the ΔCFI, ΔMc, and ΔGamma^ and significance levels and interpretation all steps of analysis are described in Table15. The acceptable levels of model fit for this study were adopted from Hu and Bentler (1999). The values adopted for assessing change in model fit, along with the equations for calculating change in model fit were adopted from Cheung & Rensvold (2002).

Table 15

Interpretation of MGCFA Significance Tests

Index	Critical Value	Equation	Interpretation
	Step1: Fit of *TAGG* Structure on Groups (Research Question 1: Are the same items associated with the same constructs for the individual group and the *TAGG* full sample?)		
RMSEA	$< .08$		If the fit meets the critical acceptability level, the same items are associated with same constructs for full *TAGG* group and comparison group, and Step 2 may be completed to assess the equality of the factor loadings for the full *TAGG* sample when they are applied to the comparison group.
SRMR	$< .08$		
SRMR	$> .95$		
NNFI	$> .95$		
	Step2: Equality of Item-Construct Relations (Research Question 2: Are the factor loadings for the individual groups significantly different than for the full *TAGG* group?)		
ΔCFI	$< -.01$	$\Delta CFI = CFI_{Step2} - CFI_{Step\ 1}$	If the fit changes by less than the critical value for at least two alternative fit indices, the factor loadings are the same for the comparison group as for the full *TAGG* sample, and model fit should be assessed.
ΔMc	$< -.02$	$\Delta Mc = Mc_{Step2} - MC_{Step\ 1}$	
ΔGamma^	$< -.001$	Δ Gamma^ = Gamma^ $_{Step\ 2}$ – Gamma^ $_{Step\ 1}$	If the fit changes by greater than the critical value for at least two of the alternative fit indices, the factor loadings may not be the same for the comparison group as for the full *TAGG* sample, and model fit may be assessed, with caution.

Table 15

Interpretation of MGCFA Significance Tests, continued

RMSEA SRMR CFI NNFI	< .08 < .08 > .95 > .95	If the change in alternative fit indices are not significant for different factor loadings for the comparison group, and the overall model fit meet critically acceptable levels, the same goals may be used regardless of degree of participation in general education. If the change in alternative fit indices are not significant for different factor loadings for the comparison group, but the overall model fit does not meet critically acceptable levels, goals may need to be differentiated based on degree of participation in general education. If the change in alternative fit indices are significant for different factor loadings for the comparison group, but the overall model fit meets critically acceptable levels, goals may need to be differentiated based on degree of participation in general education. If the change in alternative fit indices are significant for different factor loadings for the comparison group, and the overall model fit does not meet critically acceptable levels, goals may need to be differentiated based on degree of participation in general education.

The first step of the analysis tested the fit of the *TAGG* factor structure on each of the individual groups. Typical model fit statistics were used to assess whether the same items are associated with the same constructs for both the full sample and the individual group of students. This analysis also creates the fit indices necessary for the second step of the analysis. The second step of the analysis assessed for an improvement or deterioration of the model fit when the comparison group was forced to have the same factor structure as the full *TAGG* sample. This step considered whether there was a significant deterioration of the model fit and an acceptable overall fit for the more restricted model.

A Priori **Assessment of Group Overlap**

Before analyzing the data, I wanted to understand whether any of the grouping variables may be related to the participation in general education variable. For example, I was concerned that the students in the high participation in general education groups might be the same students as in the SLD or OHI groups. Should this be the case, there would be no need to analyze students in those disability categories as a separate group. Similarly, I wanted to assess whether students in some disability category may have had more exposure to concepts in transition education than students in other disability categories, again, to ensure that I was not testing two groups and simply calling them by a different name.

Tables 16-27 present the overlap between groups based on the percent of time in general education variable. Although an educator or family member complete the *TAGG-P* and *TAGG-F*, all versions of the *TAGG* are designed to assess the student. Therefore, these tables reflect students' participation in general education, disability categories, grade levels, and exposure to transition education.

Table 16

TAGG-P Student Percent of time in General Education by Grade

	Percent of Day in General Education							
Grade	< 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	Total
9th	9.7% (12)	4.8 (6)	8.1 (10)	8.1 (10)	19.4 (10)	28.2 (24)	21.8 (27)	124
10th	10.5 (17)	5.6 (9)	12.4 (20)	8.0 (13)	26.5 (43)	19.1 (31)	17.9 (29)	162
11th	7.6 (35)	8.3 (38)	9.8 (45)	13.9 (64)	24.4 (135)	17.1 (81)	18.5 (85)	480
12th	7.6 (48)	8.4 (53)	9.7 (61)	12.4. (78)	19.1 (120)	19.0 (119)	23.7 (149)	628
18-21	36.5% (27)	10.8 (8)	5.4 (4)	0.0 (0)	4.1 (3)	2.7 (2)	40.5 (30)	74
Total	139	114	140	165	302	268	320	1448

Note. Grade was not indicated for four students.

Table 17

TAGG-P Student Percent of time in General Education by Disability

	Percent of Day in General Education							
Disability	< 24.9	Total	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	Total
Autism	19.7% (12)	14.8 (9)	8.2 (5)	3.3 (2)	16.4 (10)	9.8 (6)	27.9 (17)	61
ED	25.3 (20)	5.1 (4)	10.1 (8)	6.3 (5)	11.4 (9)	19.0 (15)	22.9 (18)	79
ID	21.8 (38)	21.3 (37)	17.2 (30)	11.5 (20)	10.3 (18)	5.8 (10)	12.1 (21)	174
OHI	10.5 (21)	8.0 (16)	7.0 (14)	13.5 (27)	21.0 (42)	18.5 (37)	21.5 (43)	200
SLD	3.9 (32)	5.1 (42)	5.6 (71)	12.4 (103)	24.9 (206)	22.2 (184)	23.0 (190)	828
Total	123	108	128	157	285	252	289	1342

Note. One hundred ten student participants were identified in one of the eight other IDEA categories of disability and 17 students had no disability indicated.

Table 18

TAGG-P Student Percent of time in General Education by Transition Education

Transition Education	Percent of Day in General Education < 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	Total
Little to No Transition Education	60.5% (46)	41.5 (22)	44.9 (31)	58.9 (53)	63.2 (98)	54.6 (83)	50.0 (73)	406
Some Transition Education	39.5 (30)	58.5 (31)	55.1 (38)	41.1 (37)	36.8 (57)	45.4 (69)	50.0 (73)	355
Total	76	53	69	90	155	152	146	741

Note. Educators did not respond to items related to transition education concepts or did not know whether students had instruction in transition education for 711 students.

Table 19

TAGG-P Student Disability by Transition Education

Transition Education	Disability Category Autism	ED	ID	OHI	SLD	Total
Little to No Transition Education	41.2% (14)	56.8 (25)	45.2 (42)	52.2 (60)	60.1 (249)	402
Some Transition Education	58.8 (20)	43.2 (19)	54.8 (51)	47.8 (55)	39.9 (165)	328
Total	34	44	93	115	414	730

Note. Educators did not respond to items related to transition education concepts or did not know whether students had instruction in transition education for 646 students. Disability category was not indicated for 10 students. Eighty-nine students were identified in one of the eight IDEA categories not listed.

Table 20

TAGG-F Student Percent of time in General Education by Grade

	Percent of Day in General Education							
Grade	< 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	Total
9th	8.6%	8.7	6.7	7.6	19.1	28.6	23.8	105
	(9)	(6)	(7)	(8)	(20)	(30)	(25)	
10th	11.5	4.9	10.7	8.2	25.4	21.3	18.0	122
	(14)	(6)	(13)	(10)	(31)	(26)	(22)	
11th	8.0	8.7	9.0	9.7	25.6	20.8	18.3	289
	(23)	(25)	(26)	(28)	(74)	(60)	(53)	
12th	7.1	8.7	12.5	12.5	17.84	17.0	24.5	241
	(17)	(21)	(30)	(30)	(43)	(41)	(59)	
Total	63	58	76	76	168	157	159	757

Note. Grade was not indicated for 2 students. Thirty-four students were in programs for 18-21 year olds.

Table 21

TAGG-F Student Percent of time in General Education by Disability

	Percent of Day in General Education							
Disability	< 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	Total
ID	21.7%	20.6	16.6	13.4	8.3	5.2	11.3	97
	(n = 21)	(20)	(19)	(13)	(8)	(5)	(11)	
OHI	10.1	7.3	3.7	12.8	17.4	22.0	26.6	109
	(11)	(8)	(4)	(14)	(19)	(24)	(29)	
SLD	4.1	5.4	8.6	9.7	27.2	23.8	21.3	442
	(18)	(24)	(38)	(43)	(120)	(105)	(94)	
Total	50	52	61	70	147	134	134	648

Note. One hundred forty-five student participants were identified in one of the 10 other IDEA categories of disability.

Table 22

TAGG-F Student Percent of time in General Education by Transition Education

Transition Education	Percent of Day in General Education							Total
	< 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	
Little to No Transition Education	78.4% (n = 29)	55.6 (16)	41.4 (15)	62.5 (29)	73.8 (60)	70.0 (54)	65.7 (49)	252
Some Transition Education	21.6 (9)	44.4 (14)	58.6 (18)	37.5 (16)	26.3 (24)	30.0 (28)	34.3 (32)	141
Total	38	30	33	45	84	82	81	393

Note. Educators did not respond to items related to transition education concepts or did not know whether students had instruction in transition education for 400 students whose families participated in *TAGG* development.

Table 23

TAGG-F Student Disability by Transition Education

Transition Education	Disability Category			Total
	ID	OHI	SLD	
Little to No Transition Education	54.2% (13)	54.8 (34)	70.7 (152)	199
Some Transition Education	45.8 (11)	45.2 (28)	29.3 (63)	102
Total	24	62	215	301

Note. Educators did not respond to items related to transition education concepts or did not know whether students had instruction in transition education for 499 students. One hundred forty-five student participants were identified in one of the 10 other IDEA categories of disability.

Table 24

TAGG-S Student Percent of time in General Education by Grade

	Percent of Day in General Education							
Grade	< 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	Total
9th	10.6%	4.3	8.6	8.6	18.8	21.2	21.4	
	(n = 12)	(5)	(10)	(10)	(22)	(33)	(25)	117
10th	10.5	5.6	11.8	8.6	25.7	19.1	19.1	
	(16)	(8)	(18)	(13)	(39)	(29)	(29)	152
11th	6.5	8.2	10.0	14.2	25.2	18.0	18.0	
	(28)	(35)	(43)	(61)	(108)	(77)	(77)	429
12th	8.0	8.0	9.5	12.4	18.6	19.4	24.0	
	(47)	(47)	(56)	(73)	(109)	(114)	(141)	587
18-21	30.0	11.9	6.0	0.0	4.48	3.0	45.0	
	(20)	(8)	(4)	(0)	(3)	(2)	(30)	67
Total	123	103	131	157	281	255	302	1352

Note. Grade was not indicated for four students.

Table 25

TAGG-S Student Percent of time in General Education by Disability

	Percent of Day in General Education							
Grade	< 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	Total
Autism	14.6%	16.4	7.3	3.6	18.2	10.9	29.1	55
	(n = 8)	(9)	(4)	(2)	(10)	(6)	(16)	
ED	26.8	5.6	11.3	5.6	11.3	19.7	19.7	71
	(19)	(4)	(8)	(4)	(8)	(14)	(14)	
ID	19.5	20.8	16.8	11.7	9.7	5.8	13.6	154
	(30)	(32)	(29)	(18)	(15)	(9)	(21)	
OHI	10.6	7.9	6.7	14.3	20.1	17.5	22.8	189
	(20)	(15)	(13)	(27)	(38)	(33)	(43)	
SLD	18.6	7.9	6.9	14.3	14.6	13.2	13.3	788
	(30)	(40)	(66)	(99)	(196)	(178)	(179)	
Total	107	100	120	150	267	240	273	1257

Note. Eighty-nine student participants were identified in one of the eight other IDEA categories of disability and no disability was given for 10 students

Table 26

TAGG-S Student Percent of time in General Education by Transition Education

Transition Education	Percent of Day in General Education							Total
	< 24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7 - 100	
Little to No Transition Education	60.5% (n = 45)	41.5% (22)	44.1% (30)	58.9% (50)	63.2% (97)	55.0% (81)	50.3% (70)	395
Some Transition Education	39.5 (26)	58.5 (29)	55.9 (36)	41.1 (34)	36.8 (53)	45.0 (65)	49.7 (72)	315
Total	71	51	66	84	150	146	142	710

Note. Educators did not respond to items related to transition education concepts or did not know whether students had instruction in transition education for 646 students.

Table 27

TAGG-S Student Disability by Transition Education

Transition Education	Disability Category					Total
	Autism	ED	ID	OHI	SLD	
Little to No Transition Education	41.2% (n = 14)	58.1 (25)	44.8 (39)	52.2 (60)	60.0 (247)	397
Some Transition Education	58.8 (20)	41.9 (18)	55.2 (48)	47.8 (55)	40.0 (165)	323
Total	34	43	87	115	412	691

Note. Educators did not respond to items related to transition education concepts or did not know whether students had instruction in transition education for 646 students. Disability category was not indicated for 10 students. Eighty-nine students were identified in one of the eight IDEA categories not listed.

Based on Tables 16-27, most of the students for whom *TAGG-P*s, *TAGG-F*s, and *TAGG-S*s were completed spend more than about 70% of their day in general education. This is to be expected since data collection targeted students with mild-moderate disabilities. In reviewing the relation between disability category and participation in general education (tables 17, 21, and 25), it appears that across most disability categories, most *TAGG* student participants either spend less than a quarter of their day in general education or more than 70% of their day in general education. The only group that did not follow this pattern was the group of student participants who had intellectual disabilities. That group of students generally spent less than two thirds of their day in general education. When exposure to transition education was reviewed, it appeared that students in this study who spent about a quarter to half of their school day in general education had the most transition education, and that students with traditionally mild-moderate disabilities (ED, OHI, and SLD) had less transition education than students in this study having traditionally more severe disabilities (ID and autism).

Overall, the students who participated in the *TAGG* study appear to be spread fairly evenly across the grouping variables. Among the *TAGG* participants, there does not appear to be a strong relation between disability category and participating in general education or exposure to transition education concepts, nor does there appear to be a relation between grade level and participation in general education.

Percent of Time in General Education

One purpose of this study was to understand whether the behaviors included on the *TAGG* assessment were consistent with the behaviors families and students believe are related to the factors the *TAGG* developers have identified to be positively related to postsecondary education and competitive employment (e.g. Do families, students, and educators all believe that attaining a transition goal is a reflection of goal attainment ability?). This study also sought to further understand the theory proposed by the *TAGG* developers, particularly, whether there is a functional relation between the suitability of the family, student, and professional models across conditions that will expose them to the behaviors included on the *TAGG* (Table 8 at the end of Chapter Two). The primary variable being considered in this study is the percent of time a student spends in general education because the degree to which a student participates in general education contributes to the opportunities students may have to learn and practice the *TAGG* identified behaviors.

Percent of time in general education group participants. Tables 28-33 present the demographics for students, families, and educators whose *TAGG* assessments could be included in this portion of the study. Although this study takes the perspective that participation in general education is likely a function of level of disability and need for support, only two disability categories displayed this pattern of placement. Among students with specific learning disabilities, the number of students in each group increased as the percent of time increased. Conversely, as the percent of time in general education increased, fewer students with intellectual disabilities participated in general education.

Among teachers, there did not appear to be a pattern or relation between any of the demographic variables and the number of students who were participating in general education at different levels.

Table 28

Educator Participants for Percent of School Day in General Education Focus Group

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
n	139	114	140	165	302	268	320
Age							
18-24	2%	2	8	9	4	2	3
25-29	6	4	4	7	4	6	6
30-34	7	14	9	9	8	3	12
35-39	5	13	8	8	8	13	8
40-44	23	6	22	5	15	33	17
45-49	7	13	4	6	10	12	4
50-54	13	9	8	10	20	20	25
55-59	19	17	29	42	26	20	16
> 60	18	23	8	4	4	3	9
Mean	45.7	46.2	45.5	47.3	45.8	43.9	44.9
SD	11.0	11.8	10.7	11.2	9.7	9.5	10.5
Gender							
Male	5%	4	2	2	5	8	4
Female	95	96	98	98	95	92	96
Level of Education							
Bachelor's Degree	9%	11	9	10	8	13	7
Some Master's Coursework	35	20	27	50	21	21	14
Master's Degree	53	55	50	38	58	52	61
Educational Specialist	1	2	5	2	5	10	13
Some Doctoral Coursework	1	5	10	< 1	8	3	5
Ph.D. or Ed.D.	0	6	0	0	0	0	0
Years teaching students with disabilities							
< 3	8%	10	9	13	4	11	7
3-10	32	35	40	26	35	33	40
> 10	60	55	50	61	61	56	53
Mean	14.6	14.1	15.1	17.5	16.1	14.1	14.0
SD	9.6	10.0	11.4	12.3	9.8	9.7	10.5

Table 28

Educator Participants for Percent of School Day in General Education Focus Group, continued

Characteristic		Standard Deviation Group						
		0	1	2	3	4	5	6
Years teaching at current school								
< 3		15%	35	21	22	18	29	26
3-10		50	51	56	53	48	43	55
> 10		35	14	23	25	34	28	20
	Mean	8.6	6.7	7.8	8.9	9.9	8.6	6.7
	SD	5.7	6.2	6.7	6.6	7.9	7.6	6.5
Years teaching in current district								
< 3		44%	45	56	78	51	57	43
3-10		28	30	35	15	34	30	31
> 10		28	25	9	7	15	13	26
	Mean	12.0	10.2	11.1	12.6	12.2	10.1	9.2
	SD	8.0	8.4	9.6	9.7	9.3	9.0	8.3
Lives in district		61%	49	63	68	60	49	58
Number of students on caseload								
< 5		14%	7	4	4	9	12	3
5-10		2	23	2	3	4	10	2
10-15		29	12	25	17	17	15	13
> 15		55	58	70	76	68	63	82
	Mean	25.8	27.4	33.3	25.2	35.8	26.6	37.1
	SD	42.3	42.1	53.4	23.7	66.5	31.3	53.5
Number of college courses in transition education								
< 3		59%	78	61	76	69	73	67
3-10		23	20	36	19	25	17	28
> 10		18	2	3	5	6	10	5

Table 28

Educator Participants for Percent of School Day in General Education Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Number of professional development days in transition education							
< 10	54%	64	66	52	66	72	63
10-20	20	20	20	15	16	13	19
> 20	26	16	14	33	18	15	18
Ethnicity							
White or Caucasian	77%	78	58	75	70	83	69
Black or African-American	17	7	30	17	22	4	17
American Indian or Alaska Native	6	9	10	5	3	7	11
Mexican, Mexican-American, or Chicano	0	4	1	0	0	0	< 1
Hispanic or Latino	< 1	< 1	< 1	2	5	3	3
Asian	0	2	< 1	< 1	< 1	3	0

Table 29

Educator Participants for Percent of School Day in General Education Non Focus Group

Characteristic		Standard Deviation Group						
		0	1	2	3	4	5	6
	n	1313	1338	1312	1287	1150	1184	1132
Age								
18-24		4%	4	4	3	4	5	4
25-29		5	6	6	5	6	5	5
30-34		9	8	8	9	9	10	8
35-39		9	9	9	9	9	8	9
40-44		16	17	16	18	17	15	16
45-49		8	8	8	8	7	7	9
50-54		18	18	18	18	16	17	15
55-59		24	24	23	21	23	24	56
> 60		7	7	8	9	9	9	8
	Mean	45.3	45.4	45.2	45.1	45.3	48.8	45.6
	SD	10.3	10.4	10.3	10.3	10.6	10.6	10.4
Gender								
Male		5%	5	5	5	5	4	5
Female		95	95	95	95	95	96	95
Level of Education								
Bachelor’s Degree		9%	9	10	9	10	9	10
Some Master’s Coursework		24	25	25	21	26	25	28
Master’s Degree		54	54	54	56	53	54	52
Educational Specialist		7	7	7	7	7	6	5
Some Doctoral Coursework		6	5	5	6	4	5	5
Ph.D. or Ed.D.		< 1	0	< 1	< 1	< 1	< 1	< 1
Years teaching students with disabilities								
< 3		8%	8	8	8	9	8	23
3-10		35	35	34	36	35	35	49
> 10		56	57	58	56	56	57	28
	Mean	15.1	15.1	15.0	14.7	14.7	15.2	15.3
	SD	10.4	10.4	10.3	10.1	10.5	10.5	10.4

Table 29

Educator Participants for Percent of School Day in General Education Non Focus Group, continued

Characteristic		Standard Deviation Group						
		0	1	2	3	4	5	6
Years teaching at current school								
< 3		25%	23	24	24	25	22	23
3-10		50	50	50	50	51	52	49
> 10		25	27	26	26	24	26	28
	Mean	8.2	8.4	8.3	8.2	7.8	8.2	8.7
	SD	7.2	7.1	7.1	7.1	6.7	6.9	7.1
Years teaching in current district								
< 3		53%	53	52	49	53	51	55
3-10		30	30	59	32	29	30	29
> 10		17	17	19	19	19	19	15
	Mean	10.8	11.0	10.9	10.7	10.6	11.1	11.4
	SD	9.1	9.1	8.9	8.9	8.9	9.0	9.1
Lives in district		58%	59	58	57	58	60	58
Number of students on caseload								
< 5		7%	8	8	8	7	7	9
5-10		6	4	6	6	6	5	7
10-15		16	18	17	18	17	18	19
> 15		71	70	70	69	69	71	66
	Mean	31.9	31.6	31	32	30.1	32.2	29.5
	SD	49.8	49.5	48.7	51.6	43.6	56.3	47.8
Number of college courses in transition education								
< 3		70%	68	69	68	69	68	69
3-10		25	25	23	25	24	26	32
> 10		6	8	8	7	7	7	8

Table 29

Educator Participants for Percent of School Day in General Education Non Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Number of professional development days in transition education							
< 10	65%	64	63	65	63	62	64
10-20	17	17	17	17	17	18	17
> 20	19	20	20	18	20	21	20
Ethnicity							
White or Caucasian	73%	73	75	75	74	71	74
Black or African-American	16	17	15	15	15	19	16
American Indian or Alaska Native	7	7	7	7	8	7	6
Mexican, Mexican-American, or Chicano	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Hispanic or Latino	3	3	3	3	2	3	3
Asian	1	< 1	1	< 1	1	< 1	1

Table 30

Family Participants for Percent of School Day in General Education Focus Group

	Standard Deviation Group						
Characteristic	0	1	2	3	4	5	6
n	63	58	76	76	168	157	159
Age							
<18	0%	2	0	1	<1	<1	<1
18-24	4	2	3	0	2	3	1
25-29	1	0	0	0	0	0	0
30-34	3	3	4	6	3	4	3
35-39	14	22	25	30	22	21	15
40-44	23	28	24	33	38	24	31
45-49	17	12	10	11	25	25	24
50-54	17	14	11	11	9	5	15
55-59	10	9	13	3	4	9	4
> 60	9	10	11	4	5	3	6
Mean	46.4	44.9	45.4	42.7	43.4	43.1	45
SD	10.5	9.4	10.5	7.7	9.6	9	8.6
Primary Language							
English	96%	97	97	95	94	97	99
Other	4	3	3	5	6	3	1
Relationship to Student							
Father	18%	6	14	8	12	14	13
Grandfather	0	2	0	3	<1	0	2
Brother	1	0	0	0	0	<1	0
Male Guardian	3	0	0	1	<1	0	0
Other	1	5	3	1	<1	<1	<1
Mother	65	82	65	80	80	76	76
Grandmother	4	5	10	1	4	3	4
Aunt	0	0	0	0	1	<1	<1
Sister	1	0	1	0	<1	0	<1
Female Guardian	3	2	5	1	0	2	<1
Stepfather	1	0	0	1	0	0	<1
Stepmother	1	0	1	3	0	<1	<1

Table 30

Family Participants for Percent of School Day in General Education Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Education							
Less than high school	17%	18	20	22	16	6	8
High school diploma/GED	36	34	42	36	43	49	36
Vocational certificate	17	9	9	13	13	13	9
Associate's degree	13	16	13	13	13	14	18
Bachelor's degree	10	11	7	14	14	15	21
Master's degree	4	5	9	1	1	2	7
Doctorate/Professional degree	3	6	0	1	0	0	1
Employment							
Employed full-time	38%	52	38	51	56	55	55
Employed part-time	4	6	12	7	7	13	10
Self-employed full-time	14	3	4	1	3	3	5
Self-employed part-time	3	5	1	0	<1	5	3
Not working	30	28	31	32	22	19	20
Retired	6	5	9	5	10	3	5
Permanently disabled, not working for pay	5	2	5	4	1	3	3
Received help completing *TAGG*	8%	14	8	9	4	8	5
Help received							
Reading	50%	33	0	20	50	22	75
Writing	17	0	0	0	50	11	0
Translating	0	0	0	40	0	33	25
Explanation	17	50	100	40	0	33	25
Other	17	17	0	0	0	11	0
Ethnicity							
White	67%	45	57	69	67	67	69
Black or African-American	9	15	16	9	9	8	7
American Indian or Alaska Native	9	8	10	4	6	13	12

Table 30

Family Participants for Percent of School Day in General Education Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Mexican, Mexican American Chicano	3	11	6	7	6	4	1
Puerto Rican	0	4	0	1	0	1	0
Cuban	0	0	0	0	0	0	1
Hispanic or Latino	8	8	7	9	10	4	7
Native Hawaiian	0	0	0	0	0	0	<1
Asian	2	4	0	0	0	<1	0
Other	2	<1	3	0	<1	<1	<1
Egyptian German Spanish	0	0	0	0	<1	0	0
Middle Eastern	0	<1	0	0	0	0	0
Multiethnic	0	<1	0	0	0	0	0

Table 31

Family Participants for Percent of School Day in General Education Non Focus Group

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
n	730	735	717	717	625	636	634
Age							
<18	9%	8	8	8	9	8	7
18-24	2	2	2	2	2	2	2
25-29	< 1	1	1	1	1	< 1	1
30-34	4	4	4	3	4	3	4
35-39	21	20	20	20	19	20	22
40-44	25	25	25	24	25	25	24
45-49	19	20	20	20	17	18	18
50-54	10	11	11	11	11	13	10
55-59	6	6	5	6	6	5	7
> 60	5	5	5	6	6	7	6
Mean	45.4	45.3	45.4	45.2	45.3	45.6	45.6
SD	10.4	10.3	10.4	10.3	10.6	10.4	10.4

Table 31

Family Participants for Percent of School Day in General Education Non Focus Group, continued

Characteristic	Standard Deviation Group						
	0	1	2	3	4	5	6
Primary Language							
English	97%	96	96	97	97	97	96
Other	3	4	4	3	3	3	4
Relationship to Student							
Father	12%	14	13	14	13	13	13
Grandfather	1	< 1	1	< 1	1	1	< 1
Brother	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Male Guardian	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Other	1	1	1	1	1	1	1
Mother	77	76	77	75	75	76	76
Grandmother	4	4	4	5	5	5	5
Aunt	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Sister	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Female Guardian	2	2	1	2	2	1	2
Stepfather	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Stepmother	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Education							
Less than high school	14%	14	13	14	13	16	15
High school diploma/GED	42	42	41	42	41	40	43
Vocational certificate	12	12	13	12	12	12	13
Associate's degree	14	14	15	14	15	14	14
Bachelor's degree	14	14	15	14	14	14	12
Master's degree	4	4	3	4	4	4	3
Doctorate/Professional degree	< 1	< 1	1	< 1	1	1	1

Table 31

Family Participants for Percent of School Day in General Education Non Focus Group, continued

	Standard Deviation Group						
Characteristic	0	1	2	3	4	5	6
Employment							
Employed full-time	53%	52	53	52	51	51	51
Employed part-time	9	9	8	8	8	8	8
Self-employed full-time	3	3	4	4	4	4	4
Self-employed part-time	3	3	3	3	3	2	3
Not working	23	23	23	22	24	25	25
Retired	7	7	6	7	5	7	7
Permanently disabled, not working for pay	< 1	3	3	3	4	3	3
Received help completing *TAGG*	7%	6	7	6	8	7	7
Help received							
Reading	32%	32	35	32	31	37	26
Writing	7	8	7	7	4	5	7
Translating	14	16	14	14	14	9	15
Explanation	39	36	34	39	41	41	41
Other	7	8	10	7	10	9	11
Ethnicity							
White	65%	67	67	65	65	65	65
Black or African-American	10	9	10	11	11	10	13
American Indian or Alaska Native	10	8	10	10	9	10	9
Mexican, Mexican American Chicano	5	4	5	5	5	5	6
Puerto Rican	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Cuban	< 1	< 1	< 1	< 1	< 1	< 1	0
Hispanic or Latino	7	6	7	7	8	7	7
Native Hawaiian	0	0	0	<1	< 1	0	0
Asian	1	1	1	2	2	1	2
Other	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Egyptian	< 1	< 1	< 1	< 1	< 1	< 1	0
German	0	< 1	< 1	< 1	< 1	< 1	< 1
Spanish	< 1	0	< 1	< 1	< 1	< 1	< 1
Middle Eastern	< 1	0	< 1	< 1	< 1	< 1	< 1
Multiethnic	< 1	< 1	< 1	< 1	< 1	< 1	< 1

Table 32

Student Participants for Percent of School Day in General Education Focus Group

Characteristic	Standard Deviation Group						
	0	1	2	3	4	5	6
n	123	103	131	157	281	255	302
Age							
14	4%	< 1	1	2	1	5	2
15	6	3	7	4	7	8	8
16	10	11	11	13	16	14	9
17	18	28	25	23	32	27	21
18-21	62	57	55	58	46	46	60
Mean	17.7	17.6	17.4	17.4	17.1	17.1	17.6
SD	1.6	1.1	1.3	1.1	1.5	21.3	1.4
Gender							
Male	60%	64	59	57	62	60	56
Female	40	36	41	43	39	40	44
Primary Disability							
Autism	9%	8	4	1	3	2	5
Deaf-Blindness	0	0	0	0	< 1	< 1	0
Emotional Disturbance	15	4	6	3	3	6	6
Hearing Impaired	< 1	0	< 1	1	2	1	< 1
Intellectual Disability	28	33	22	12	6	4	7
Multiple Disability	3	< 1	0	0	< 1	0	3
Orthopedic Impairment	0	< 1	2	0	< 1	< 1	1
Other Health Impairment	15	14	10	17	14	14	14
Specific Learning Disability	23	38	51	63	68	68	60
Speech/Language Disability	< 1	2	1	1	1	3	1
Traumatic Brain Injury	0	0	1	< 1	0	< 1	2
Visual Impairment/Blindness	0	0	< 1	0	< 1	< 1	< 1
Other	6	0	1	< 1	< 1	0	< 1

Table 32

Student Participants for Percent of School Day in General Education Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Level of Disability							
Mild-Moderate	87%	89	83	91	90	97	88
Severe-Profound	9	8	5	7	3	< 1	4
Not Indicated	4	3	11	1	8	2	8
Attendance							
Absent ≤ 1 day/month	58%	59	55	57	62	60	62
Absent 2-3 days/month	32	32	30	30	27	30	25
Absent ≥ 4 days/month	11	8	16	13	11	9	13
Attended IEP meeting	90%	92	87	88	89	86	88
Actively participated in IEP meeting	54	58	64	66	62	64	65
Led IEP meeting	12	11	8	8	12	7	11
Years teacher has known student							
< 1 year	38%	34	47	41	26	30	36
1-2 years	23	31	28	25	38	26	28
2-3 years	24	20	14	15	17	33	26
3-4 years	10	6	1	3	7	2	3
4-5 years	3	9	8	12	9	5	3
> 5 years	3	0	1	4	4	4	4
Mean	2.3	2.5	2.2	2.6	2.9	2.8	2.5
SD	1.7	1.6	1.7	2.1	2.4	2.4	2.5
Years student has attended school							
< 1 year	21%	13	24	16	17	16	23
1-2 years	22	15	21	27	19	15	21
2-3 years	18	19	21	24	28	22	17
3-4 years	21	30	23	13	15	33	28
4-5 years	18	24	11	21	21	13	11
> 5 years	12	18	9	21	17	12	10
Mean	2.9	3.5	2.9	3.4	3.4	3.2	3.1
SD	2.1	2.0	2.2	2.8	2.8	2.3	2.8

Table 32

Student Participants for Percent of School Day in General Education Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Years student has been enrolled in district							
< 1 year	21%	12	17	12	12	8	10
1-2 years	10	3	2	3	6	7	6
2-3 years	5	5	9	3	4	8	5
3-4 years	2	2	< 1	3	4	3	1
4-5 years	23	20	16	27	24	30	26
> 5 years	39	57	56	52	50	45	52
Mean	6.7	8.9	7.9	7.8	7.8	7.6	8.3
SD	4.8	4.4	4.9	4.4	4.3	4.1	4.7
Eligible for free/reduced lunch	61%	67	65	64	59	52	51
Receives free/reduced lunch	61	68	67	64	58	52	53
Receives ESL services	3	5	2	6	3	2	6

Table 32

Student Participants for Percent of School Day in General Education Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Ethnicity							
American	0%	0	0	0	0	0	< 1
American Indian or Alaska Native	7	9	7	5	7	13	15
Arabic	0	0	0	0	0	0	< 1
Asian	0	2	1	0	1	< 1	0
Black or African-American	15	35	31	18	20	14	17
Cuban	0	0	0	0	< 1	< 1	0
Egyptian	0	0	0	0	0	0	< 1
Guamanian, Chamorro	0	0	0	0	0	< 1	0
Hispanic or Latino	5	9	7	11	10	7	9
Irish	0	0	0	0	< 1	0	0
Multiethnic	0	< 1	0	< 1	0	< 1	0
Mexican, Mexican-American, or Chicano	9	7	7	7	8	5	5
Native Hawaiian	0	0	0	0	0	0	1
Other (Not Indicated)	< 1	0	0	< 1	< 1	0	< 1
Puerto Rican	< 1	0	0	< 1	0	< 1	< 1
White or Caucasian	63	38	48	56	51	57	52

Table 33

Student Participants for Percent of School Day in General Education Non Focus Group

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
n	1273	1293	1265	1239	1115	1141	1094
Age							
14	2%	2	2	2	3	2	3
15	7	7	7	7	7	7	7
16	12	12	12	12	11	12	13
17	26	25	26	26	24	25	27
18-21	53	54	53	53	52	54	50
Mean	17.4	17.4	17.4	17.4	17.4	17.5	17.4
SD	1.2	1.3	1.3	1.3	1.3	1.3	1.2
Gender							
Male	59%	58	59	59	58	58	60
Female	41	42	41	41	42	42	40
Primary Disability							
Autism	4%	4	4	5	4	5	4
Deaf-Blindness	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Emotional Disturbance	4	5	5	6	6	5	5
Hearing Impaired	1	1	1	1	1	< 1	1
Intellectual Disability	10	10	10	12	13	13	13
Multiple Disability	< 1	1	1	1	1	1	< 1
Orthopedic Impairment	< 1	1	1	1	1	1	1
Other Health Impairment	< 1	< 1	< 1	< 1	< 1	< 1	< 1
Specific Learning Disability	14	14	15	14	14	14	14
Speech/Language Disability	62	60	59	58	56	56	58
Traumatic Brain Injury	2	2	2	2	2	1	2
Visual Impairment/Blindness	< 1	< 1	0	< 1	< 1	< 1	< 1
Other	< 1	< 1	0	< 1	< 1	< 1	< 1
Level of Disability							
Mild-Moderate	90%	90	90	90	90	88	90
Severe-Profound	6	6	5	6	5	7	5
Not Indicated	4	4	5	4	5	5	5

Table 33

Student Participants for Percent of School Day in General Education Non Focus Group, continued

Characteristic	Standard Deviation Group 0	1	2	3	4	5	6
Attendance							
Absent ≤ 1 day/month	60%	60	60	60	60	59	59
Absent 2-3 days/month	29	29	29	29	29	29	30
Absent ≥ 4 days/month	12	12	11	11	12	12	11
Attended IEP meeting	90%	90	88	88	88	88	88
Actively participated in IEP meeting	63	63	62	62	63	62	61
Led IEP meeting	7	8	10	10	9	10	9
Years teacher has known student							
< 1 year	33%	33	31	32	35	34	33
1-2 years	30	30	30	31	28	31	31
2-3 years	23	23	24	24	25	20	22
3-4 years	4	4	4	4	3	5	4
4-5 years	7	6	6	6	6	7	7
> 5 years	6	4	4	3	3	3	3
Mean	1.5	1.5	1.5	1.5	1.4	1.4	1.5
SD	1.9	2.0	2.0	1.9	1.9	1.9	1.9
Years student has attended school							
< 1 year	19%	20	19	19	20	20	19
1-2 years	14	15	14	14	14	15	13
2-3 years	24	24	24	23	22	24	26
3-4 years	26	25	26	27	28	23	24
4-5 years	3	3	4	4	3	4	4
> 5 years	13	13	13	12	12	13	14
Mean	2.2	2.2	2.2	2.2	2.2	2.2	12.3
SD	1.6	1.6	1.6	1.6	1.6	1.6	1.6

Table 33

Student Participants for Percent of School Day in General Education Non Focus Group, continued

	Standard Deviation Group						
Characteristic	0	1	2	3	4	5	6
Years student has been enrolled in district							
< 1 year	9%	10	9	9	9	10	10
1-2 years	5	6	6	6	6	5	6
2-3 years	6	6	6	6	6	6	6
3-4 years	2	2	3	2	2	2	3
4-5 years	26	26	27	26	26	25	26
> 5 years	51	50	50	51	51	52	50
Mean	6.4	6.2	6.3	6.3	6.3	6.4	6.2
SD	3.8	3.8	3.8	3.8	3.8	3.9	3.9
Eligible for free/reduced lunch	57%	57	57	57	57	59	60
Receives free/reduced lunch	57	57	57	57	58	59	60
Receives ESL services	4	4	4	3	4	4	3
Ethnicity							
American	0%	< 1	<1	<1	<1	<1	0
American Indian or Alaska Native	7	10	10	6	9	13	9
Arabic	0	< 1	0	0	0	0	< 1
Asian	2	2	1	0	1	< 1	1
Black or African-American	20	19	19	20	21	20	21
Cuban	< 1	< 1	0	0	< 1	< 1	0
Egyptian	< 1	< 1	0	0	0	0	< 1
Guamanian, Chamorro	< 1	< 1	0	0	0	< 1	0
Hispanic or Latino	11	10	7	11	9	7	7
Irish	< 1	0	< 1	< 1	< 1	<1	<1
Multiethnic	< 1	< 1	0	< 1	0	< 1	0
Mexican, Mexican-American, or Chicano	8	8	7	8	7	5	8
Native Hawaiian	< 1	< 1	0	0	0	0	1
Other (Not Indicated)	< 1	0	0	< 1	< 1	0	< 1
Puerto Rican	< 1	< 1	0	< 1	0	< 1	< 1
White or Caucasian	51	53	53	51	53	51	52

Transition Education, Grade Level, and Categorical Label Groups

Grouping students based on the percent of time they spent in general education was a novel approach to understanding the consistency of a transition assessment's framework. This study acknowledges that students may learn many of the behaviors included on the *TAGG* in Transition Education courses and through new expectations students face during high school. Therefore, the study was expanded to understand whether exposure to and opportunities to practice the *TAGG* behaviors through these venues contributed to changes in beliefs about the meaning of the *TAGG* structure.

Transition Education

A key rationale for this study was that students with disabilities were more likely to learn the skills the *TAGG* assesses if they spent more time in special education, and that the appropriateness of the goals suggested by the *TAGG* would change as students spent more or less time in general education. For example, if a student has had instruction in actively participating in his IEP meeting, but still scores low on the construct, he may need goals for behaviors that are not included on the *TAGG* assessments. This rationale relied heavily on the assumption that students who did not participate in special education did not have the knowledge necessary to understand the *TAGG* items—particularly the items associated with disability awareness and involvement in the IEP.

Rationale for transition education analysis. It should be acknowledged that some students spend most of their school day in general education, and take (or have taken) a transition course in which they learned the transition-specific language used in the *TAGG* items. It is expected that these students will have experienced a degree of indoctrination of the value of the behaviors included on the *TAGG*, particularly given that the professional educators who participated in this study were recruited through transition education email lists and at transition related conferences. Therefore, the purpose of this analysis was to understand whether the educators and families of students who had some transition education, along with the students themselves, held the same beliefs about the behaviors associated with the *TAGG* constructs those who completed the *TAGG* for students who did not have transition education.

Due to the nature of the skills assessed by the *TAGG*, it was expected that the *TAGG* structure would be most appropriate for students who spend most of their time in special education (and can practice skills such as working in groups, or talking about how their disability impacts their lives), or for students who have had some transition education and spend most of the their time in general education (and have learned how to lead their IEP meetings). Students who spent most of their time in general education and had little to no transition education were expected to be impacted by language used on some *TAGG* items, and thus not associate the behaviors with the constructs.

Many of the items use special education specific language that may interfere with some students' ability to understand their intended meaning. Based on this, and because the demographics used to quantify transition education (described in the computation of transition education variable section) were directly related to *TAGG* constructs, steps needed to be taken to understand whether the *TAGG* scores were accurate for students who had not had transition education. Therefore, unlike the prior analysis, which sought to understand whether the suggested goals were appropriate across different groups of students, this analysis sought to understand whether all of the students who take the *TAGG* understand the meaning of the questions.

Computation of transition education variable. The *TAGG* demographic forms asked educators to respond to questions addressing whether students had instruction to learn to be actively involved in their IEP meetings, to lead their IEP meetings, and around disability awareness. Educators could respond by indicating yes, no, or don't know. Students were then divided into groups based on the level of transition education they had, as described in Table 34.

Table 34

Definitions of Level of Transition Education

	Little to No Transition Education	Some Transition Education
Phase I	No instruction in leading IEP *OR* No disability awareness instruction	Instruction in leading IEP *AND* Disability awareness instruction
Phases II and III	No instruction in IEP meeting participation *OR* No instruction in leading IEP *AND* No disability awareness instruction	Instruction in IEP meeting participation *AND/OR* Instruction in leading IEP *AND* Disability awareness instruction

Phase I level of transition education variable. The first phase of the *TAGG* study asked teachers to indicate whether students had instruction in how to lead their IEP meetings and whether students had disability awareness instruction. The items were coded one for yes, zero for no, and missing for “Don’t know.” If an item was missing, no transition education total could be calculated.

If an educator indicated that the student had not had instruction in how to lead their IEP meetings or the student had instruction in *either* leading his IEP meeting *or* disability awareness, the student was coded as having little to no transition education. If the educator indicated the student had both instruction in leading his IEP meeting, *and* disability awareness, he was coded as having transition education.

Phase II and III level of transition education variable. The second and third phases of the *TAGG* study asked teachers to indicate whether students had instruction in how to lead their IEP meetings, whether students had instruction in how to actively participate in their IEP meetings, and whether students had disability awareness instruction. The items were coded one for yes, zero for no, and missing for don’t know. If an item was missing, no transition education total could be calculated.

If an educator indicated that the student either had no instruction in how to lead their IEP meetings or had disability awareness instruction or had neither, the student was coded as having little to no transition education. If the educator indicated the student had two types of instruction—leading his IEP meeting *and/or* actively participating in his IEP meetings *and* disability awareness, he was coded as having transition education.

Students who participated in *TAGG* development spent an average of 67.9% of their school day in general education. Table 35 indicates the percent of students who had little to no transition or some transition education among students who participated in general education above average or below average.

Table 35

Percent of Students at Each Transition Education Level by Percent of Time in General Education

Percent of Time in General Education	Little to No Transition Education	Some Transition Education
$\leq$ 67.9%	22% (n = 176)	19% (n = 146)
> 67.9%	33 (255)	26 (203)

Note. Missing=757

Based on Table 18, it can be noted percent of time in general education and level of transition education could be calculated for about half of the sample (n = 780, 50.74%). Of those participants, 349 had disability awareness and instruction in leading or participating in their IEP meetings. Four hundred thirty-one of the participants had either instruction in disability awareness or leading/participating in their IEP meetings, or instruction in neither area. A greater proportion of students who participated in general education appear to have had some transition education, while students who participated in general education for a below average portion of their day appear to have had less transition education.

Transition education participants. Tables 36-41 provide the professional and student demographics for each transition education group, respectively. Substantially fewer students could be included in this analysis, largely due to missing information on the student demographic forms completed by the educators (see the calculation of the transition education variable section). Worth noting is the jump in level of transition education among students from age 16 to 17 (Tables 40 and 41). At the age of 16, less than 10% of students had some transition education, but by 17, nearly 25% reported having some transition education, and by 18, over 70% of students had some transition education.

Table 36

Educator Participants by Instruction in Transition Education Focus Group

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
	n	405	333
Age			
18-24		2%	5
25-29		9	6
30-34		< 1	6
35-39		16	10
40-44		10	28
45-49		13	1
50-54		17	12
55-59		30	19
> 60		3	14
	Mean	45.7	45.3
	SD	9.77	11.1
Gender			
Male		5%	5
Female		95	95
Level of Education			
Bachelor's Degree		11	7
Some Master's Coursework		30	24
Master's Degree		24	50
Educational Specialist		7	8
Some Doctoral Coursework		6	10
Ph.D. or Ed.D.		0	< 1
Years teaching students with disabilities			
< 3		12	3
3-10		31	40
> 10		57	27
	Mean	14.0	15.8
	SD	10.1	10.7

Table 36

Educator Participants by Instruction in Transition Education Focus Group, continued

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Years teaching at current school			
< 3		27	16
3-10		50	54
> 10		23	30
	Mean	7.7	8.7
	SD	6.7	6.6
Years teaching in current district			
< 3		70	36
3-10		26	36
> 10		5	28
	Mean	8.7	12.4
	SD	8.4	8.9
Lives in district		58	60
Number of students on caseload			
< 5		6	6
5-10		< 1	7
10-15		12	17
> 15		81	70
	Mean	24.0	29.1
	SD	21.0	36.9
Number of college courses in transition education			
< 3		73	62
3-10		16	31
> 10		10	7
Number of professional development days in transition education			
< 10		62	67
10-20		18	7
> 20		20	26

Table 36

Educator Participants by Instruction in Transition Education Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Ethnicity		
White or Caucasian	78	75
Black or African-American	7	16
American Indian or Alaska Native	12	7
Mexican, Mexican-American, or Chicano	0	0
Hispanic or Latino	3	< 1
Asian	0	3

Table 37

Educator Participants by Instruction in Transition Education Non Focus Group

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
	n	336	408
Age			
18-24		5%	4
25-29		4	5
30-34		12	9
35-39		6	9
40-44		19	12
45-49		6	10
50-54		17	19
55-59		21	25
> 60		10	7
	Mean	45.3	45.4
	SD	10.7	10.2
Gender			
Male		5%	5
Female		95	95

Table 37

Educator Participants by Instruction in Transition Education Non Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Level of Education		
Bachelor's Degree	9%	10
Some Master's Coursework	23	25
Master's Degree	57	55
Educational Specialist	7	6
Some Doctoral Coursework	5	4
Ph.D. or Ed.D.	< 1	< 1
Years teaching students with disabilities		
< 3	7%	10
3-10	36	34
> 10	57	57
Mean	15.4	14.8
SD	10.5	10.3
Years teaching at current school		
< 3	22%	26
3-10	50	49
> 10	27	25
Mean	8.5	8.1
SD	7.2	7.2
Years teaching in current district		
< 3	46%	57
3-10	31	58
> 10	23	15
Mean	11.4	10.5
SD	9.2	9.0
Lives in district	58%	58

Table 37

Educator Participants by Instruction in Transition Education Non Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Some Transition Education
Number of students on caseload		
< 5	9%	8
5-10	8	5
10-15	20	18
> 15	65	69
Mean	34.2	31.9
SD	56.6	52.4
Number of college courses in transition education		
< 3	68%	71
3-10	26	22
> 10	6	7
Number of professional development days in transition education		
< 10	64%	62
10-20	17	40
> 20	19	18
Ethnicity		
White or Caucasian	71%	72
Black or African-American	19	16
American Indian or Alaska Native	5	7
Mexican, Mexican-American, or Chicano	< 1	< 1
Hispanic or Latino	2	3
Asian	1	<1

Table 38

Family Participants by Instruction in Transition Education Focus Group

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
	n	252	141
Age			
<18		7%	18
18-24		1	3
25-29		1	1
30-34		3	6
35-39		24	16
40-44		26	23
45-49		16	13
50-54		10	9
55-59		6	5
> 60		6	6
	Mean	44.2	43.6
	SD	8.7	10.7
Primary Language			
English		97%	93
Other		3	7
Relationship to Student			
Father		13%	10
Grandfather		0	<1
Brother		0	0
Male Guardian		0	<1
Other		1	1
Mother		77	77
Grandmother		5	4
Aunt		<1	0
Sister		<1	<1
Female Guardian		1	3
Stepfather		0	<1
Stepmother		<1	2

Table 38

Family Participants by Instruction in Transition Education Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Education		
Less than high school	13%	17
High school diploma/GED	37	34
Vocational certificate	17	11
Associate's degree	18	16
Bachelor's degree	12	18
Master's degree	3	3
Doctorate/Professional degree	<1	1
Employment		
Employed full-time	53%	51
Employed part-time	8	9
Self-employed full-time	5	6
Self-employed part-time	2	1
Not working	24	19
Retired	8	3
Permanently disabled, not working for pay	<1	9
Received help completing *TAGG*	6%	9
Help received		
Reading	43%	33
Writing	7	8
Translating	14	17
Explanation	36	33
Other	0	8

Table 38

Family Participants by Instruction in Transition Education Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Ethnicity		
White	68%	46
Black or African-American	9	12
American Indian or Alaska Native	10	15
Mexican, Mexican-American, Chicano	5	12
Puerto Rican	1	0
Cuban	0	0
Hispanic or Latino	6	11
Native Hawaiian	0	1
Asian	<1	1
Other	0	0
Egyptian	0	0
German	0	0
Spanish	0	0
Middle Eastern	0	0
Multiethnic	0	0
Spanish	0	0
Multiethnic	<1	1

Table 39

Family Participants by Instruction in Transition Education Non Focus Group

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
	n	489	600
Age			
<18		9%	6
18-24		2	2
25-29		< 1	1
30-34		4	3
35-39		18	21
40-44		24	25
45-49		20	20
50-54		11	11
55-59		6	6
> 60		6	6
	Mean	44.3	44.0
	SD	9.4	9.8
Primary Language			
English		97%	97
Other		3	3
Relationship to Student			
Father		13%	14
Grandfather		1	1
Brother		< 1	< 1
Male Guardian		< 1	< 1
Other		1	1
Mother		76	76
Grandmother		4	5
Aunt		< 1	< 1
Sister		< 1	< 1
Female Guardian		2	2
Stepfather		< 1	< 1
Stepmother		< 1	< 1

Table 39

Family Participants by Instruction in Transition Education Non Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Education		
Less than high school	14%	13
High school diploma/GED	43	43
Vocational certificate	10	12
Associate's degree	12	14
Bachelor's degree	15	14
Master's degree	4	4
Doctorate/Professional degree	2	< 1
Employment		
Employed full-time	52%	52
Employed part-time	8	8
Self-employed full-time	4	3
Self-employed part-time	3	3
Not working	23	24
Retired	6	7
Permanently disabled, not working for pay	5	2
Received help completing *TAGG*	7%	6
Help received		
Reading	32%	32
Writing	5	5
Translating	11	11
Explanation	37	42
Other	16	11

Table 39

Family Participants by Instruction in Transition Education Non Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Ethnicity		
White	64%	68
Black or African-American	11	10
American Indian or Alaska Native	9	9
Mexican, Mexican-American, Chicano	6	4
Puerto Rican	< 1	< 1
Cuban	< 1	< 1
Hispanic or Latino	8	7
Native Hawaiian	< 1	< 1
Asian	2	2
Other	< 1	< 1
Egyptian	< 1	< 1
German	< 1	< 1
Middle Eastern	< 1	< 1
Spanish	0	< 1
Multiethnic	< 1	< 1

Table 40

Student Participants by Instruction in Transition Education Focus Group

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
	n	395	315
Age			
14		6%	0
15		13	0
16		16	7
17		23	23
18-21		34	71
	Mean	17.0	17.9
	SD	1.4	1.0
Gender			
Male		58%	61
Female		42	38
Primary Disability			
Autism		3%	6
Deaf-Blindness		< 1	< 1
Emotional Disturbance		6	6
Hearing Impaired		2	< 1
Intellectual Disability		10	15
Multiple Disability		0	1
Orthopedic Impairment		0	< 1
Other Health Impairment		14	17
Specific Learning Disability		62	51
Speech/Language Disability		< 1	1
Traumatic Brain Injury		0	< 1
Visual Impairment/Blindness		< 1	< 1
Other		< 1	0
Level of Disability			
Mild-Moderate		97%	95
Severe-Profound		3	5
Not Indicated		0	0
Attendance			
Absent ≤ 1 day/month		62%	59
Absent 2-3 days/month		28	29
Absent ≥ 4 days/month		10	12

Table 40

Student Participants by Instruction in Transition Education Focus Group, continued

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Attended IEP meeting		84%	89
Actively participated in IEP meeting		57	80
Led IEP meeting		13	14
Years teacher has known student			
< 1 year		33%	27
1-2 years		28	27
2-3 years		26	27
3-4 years		5	8
4-5 years		5	9
> 5 years		3	2
	Mean	2.5	2.8
	SD	2.2	2.0
Years student has attended school			
< 1 year		21%	15
1-2 years		17	8
2-3 years		22	20
3-4 years		20	35
4-5 years		5	7
> 5 years		17	17
	Mean	3.2	3.6
	SD	3.0	2.4
Years student has been enrolled in district			
< 1 year		10%	10
1-2 years		6	4
2-3 years		7	5
3-4 years		4	3
4-5 years		32	23
> 5 years		40	56
	Mean	7.0	8.4
	SD	4.3	4.5

Table 40

Student Participants by Instruction in Transition Education Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Eligible for free/reduced lunch	58%	58
Receives free/reduced lunch	56	57
Receives ESL services	3	6
Ethnicity		
American	< 1	0
American Indian or Alaska Native	11	14
Asian	< 1	< 1
Black or African-American	10	30
Cuban	0	< 1
Guamanian, Chamorro	0	< 1
Hispanic or Latino	12	5
Irish	0	< 1
Multiethnic	0	< 1
Mexican, Mexican-American, or Chicano	5	8
Native Hawaiian	0	< 1
Other (Not Indicated)	0	0
Puerto Rican	< 1	0
White or Caucasian	58	40

Table 41

Student Participants by Instruction in Transition Education Non Focus Group

Characteristic		Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
	n	346	426
Age			
14		7%	6
15		24	13
16		48	17
17		13	23
18-21		5	41
	Mean	17.9	16.9
	SD	1.0	1.4

Table 41

Student Participants by Instruction in Transition Education Non Focus Group, continued

Characteristic	Transition Education: Little to no Transition Education	Transition Education: Some Transition Education
Gender		
Male	60%	58
Female	40	42
Primary Disability		
Autism	6%	4
Deaf-Blindness	< 1	< 1
Emotional Disturbance	5	6
Hearing Impaired	1	2
Intellectual Disability	15	10
Multiple Disability	1	0
Orthopedic Impairment	0	< 1
Other Health Impairment	1	0
Specific Learning Disability	17	15
Speech/Language Disability	51	63
Traumatic Brain Injury	1	< 1
Visual Impairment/Blindness	< 1	0
Other	< 1	< 1
Level of Disability		
Mild-Moderate	69%	97
Severe-Profound	4	3
Not Indicated	0	0
Attendance		
Absent ≤ 1 day/month	58%	62
Absent 2-3 days/month	30	27
Absent ≥ 4 days/month	12	11

Table 41

Student Participants by Instruction in Transition Education Non Focus Group, continued

	Transition Education	
Characteristic	Little to no Transition Education	Some Transition Education
Attended IEP meeting	89%	84
Actively participated in IEP meeting	79	57
Led IEP meeting	13	13
Years teacher has known student		
< 1 year	27%	33
1-2 years	28	29
2-3 years	28	25
3-4 years	8	5
4-5 years	8	5
> 5 years	2	3
Mean	1.6	1.4
SD	1.7	1.8
Years student has attended school		
< 1 year	14%	21
1-2 years	8	17
2-3 years	20	21
3-4 years	34	20
4-5 years	7	5
> 5 years	17	16
Mean	2.6	2.1
SD	1.5	1.7
Years student has been enrolled in district		
< 1 year	7%	11
1-2 years	5	6
2-3 years	5	8
3-4 years	3	4
4-5 years	24	32
> 5 years	60	40
Mean	6.8	5.5
SD	3.8	3.8

Table 41

Student Participants by Instruction in Transition Education Non Focus Group, continued

	Transition Education	
Characteristic	Little to no Transition Education	Some Transition Education
Eligible for free/reduced lunch	58%	58
Receives free/reduced lunch	57	56
Receives ESL services	5	3
Ethnicity		
American	0%	< 1
American Indian or Alaska Native	15	12
Asian	3	2
Black or African-American	30	11
Cuban	< 1	0
Guamanian, Chamorro	0	0
Hispanic or Latino	7	13
Irish	< 1	0
Multiethnic	< 1	0
Mexican, Mexican-American, or Chicano	4	6
Native Hawaiian	0	0
Other (Not Indicated)	0	0
Puerto Rican	0	0
White or Caucasian	40	58

Data analysis of transition education groups. This analysis followed the same steps described for the standard deviation groups and examined the properties of the *TAGG-P, TAGG-F,* and *TAGG-S* across professionals, and families whose students had little-to-no or some transition education, and students who have had little-to-no or some transition education.

First, the factor structure for each version of the *TAGG* was applied to its respective groups to assess for the suitability of the factor structure and to ensure that the same items were associated with the same constructs for students who had each level of transition education. If the tested structure was suitable for the groups, the second step was applied to ensure the factor loadings were equal across groups.

Grade Level

The *TAGG* developers believe the skills the *TAGG* assesses can be learned, and that students who are just beginning high school should have lower *TAGG* scores than students who are nearly finished with high school. In order to compare scores across grades, the *TAGG* structure and the goals it generates based on that structure needed to be reviewed for suitability across grade levels. If the *TAGG* structure was not equivalent across grades, scores could not be compared from one year to the next, and grade level would need to be considered when goals are suggested as part of the student's score profile.

Rationale for grade level analysis. It is easy to understand why students who are just beginning high school would score lower on the *TAGG* than students who are exiting high school—the *TAGG* skills are, for the most part, *learned* skills, not unwavering traits. However, before the scores across grades can be compared, the *TAGG* developers needed to ensure that students in different grades use the same behaviors to define the scores the *TAGG* produces. For example, a senior in high school may have had much more exposure to the *TAGG* behaviors and skills, and may have experienced more indoctrination with the self-determination framework than a student in 9^{th} grade. Therefore, the purpose of this grouping is to understand whether *TAGG* scores vary across grade due to actual differences in students' abilities or whether the differences related to how appropriate the *TAGG* framework is for each grade.

Grade level participants. This analysis included students in grades 9-12 and in programs for students ages 18-21. Students with disabilities who participated in *TAGG* development were remarkable in that they were, on average, in the grade that corresponded with their age, with a standard deviation of less than one year for each grade. Tables 42 and 47 describe the participants who completed *TAGG* forms for students at each grade level.

Table 42

Educator Participants by Grade Focus Group

Characteristic		Grade Level			
	9th	10th	11th	12th	18-21
n	124	162	480	628	74
Age					
18-24	1%	2	5	5	0
25-29	5	4	3	7	5
30-34	2	11	10	9	3
35-39	17	10	9	8	3
40-44	4	5	22	15	30
45-49	18	12	9	6	0
50-54	28	20	14	14	45
55-59	23	30	21	26	5
> 60	2	5	9	10	8
Mean	44	44.8	44.5	46.4	46.5
SD	9.5	10.3	10.2	10.8	9.2
Gender					
Male	10%	1	3	7	0
Female	90	99	97	93	100
Level of Education					
Bachelor’s Degree	8%	6	10	11	3
Some Master’s Coursework	16	29	23	28	14
Master’s Degree	52	58	52	51	77
Educational Specialist	15	6	7	6	0
Some Doctoral Coursework	10	1	6	4	5
Ph.D. or Ed.D.	0	0	< 1	< 1	1
Years teaching students with disabilities					
< 3	9%	12	6	10	0
3-10	40	35	38	32	31
> 10	50	53	56	58	69
Mean	11.6	14.3	14.6	15.9	17.3
SD	8.7	1.0.8	10.0	10.9	8.9

Table 42

Educator Participants by Grade Focus Group, continued

Characteristic		Grade Level				
		9th	10th	11th	12th	18-21
Years teaching at current school						
< 3		37%	29	22	23	7
3-10		34	44	53	51	68
> 10		29	27	25	26	26
	Mean	7.4	8.3	8.3	8.5	7.3
	SD	7.9	7.7	7.1	6.8	5.3
Years teaching in current district						
< 3		64%	60	49	56	11
3-10		27	28	37	27	16
> 10		10	12	14	17	73
	Mean	7.7	10.3	10.8	11.4	14.9
	SD	7.6	9.2	8.8	9.4	6.3
Lives in district		51%	55	57	61	59
Number of students on caseload						
< 5		14%	11	6	6	18
5-10		6	2	7	5	14
10-15		13	18	21	17	9
> 15		67	69	67	73	60
	Mean	23.3	32.5	33.1	30.9	27.7
	SD	35.0	52.9	60.7	41.7	41.0
Number of college courses in transition education						
< 3		65%	78	72	68	49
3-10		13	13	23	28	34
> 10		22	8	5	4	17

Table 42

Educator Participants by Grade Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th	18-21
Number of professional development days in transition education	72%	60	67	59	76
< 10	12	23	15	18	18
10-20	16	17	18	23	7
> 20					
Ethnicity					
White or Caucasian	69	62	73	74	88
Black or African-American	12	20	16	16	9
American Indian or Alaska Native	12	12	6	7	0
Mexican, Mexican-American, or Chicano	5	5	3	2	0
Hispanic or Latino	0	0	0	< 1	0
Asian	0	0	2	< 1	3

Table 43

Educator Participants by Grade Non Focus Group

Characteristic	Grade Level 9th	10th	11th	12th	18-21
n	1328	1292	932	824	1378
Age					
18-24	4%	4	4	3	4
25-29	5	6	6	4	5
30-34	9	8	8	8	9
35-39	8	9	9	10	9
40-44	17	18	14	17	16
45-49	7	7	7	10	8
50-54	16	17	18	20	16
55-59	24	23	25	21	24
> 60	9	9	8	7	8
Mean	45.5	45.5	45.9	44.7	45.4
SD	10.5	10.5	10.5	10.1	10.5

Table 43

Educator Participants by Grade Non Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th	18-21
Gender					
Male	4%	4	6	3	5
Female	96	95	94	97	95
Level of Education					
Bachelor's Degree	10%	10	9	8	10
Some Master's Coursework	26	24	25	22	25
Master's Degree	54	53	54	56	52
Educational Specialist	6	7	7	8	7
Some Doctoral Coursework	5	6	5	6	5
Ph.D. or Ed.D.	< 1	< 1	< 1	< 1	< 1
Years teaching students with disabilities					
< 3	8%	8	9	7	9
3-10	34	35	34	37	35
> 10	57	57	57	56	56
Mean	15.3	15.1	15.2	14.3	14.9
SD	10.5	10.3	10.6	10.0	10.4
Years teaching at current school					
< 3	23%	23	25	24	24
3-10	52	51	49	50	49
> 10	26	26	27	26	26
Mean	8.3	8.3	8.3	8.1	8.3
SD	7.0	7.0	7.0	7.2	7.1
Years teaching in current district					
< 3	52%	52	54	50	55
3-10	30	29	26	32	30
> 10	19	19	20	18	15
Mean	11.2	11.0	11.0	10.6	10.7
SD	9.1	9.0	9.1	8.7	9.1

Table 43

Educator Participants by Grade Non Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th	18-21
Lives in district	59%	58	59	56	58
Number of students on caseload					
< 5	7%	7	9	9	7
5-10	6	6	5	6	5
10-15	18	17	16	18	18
> 15	69	69	70	67	70
Mean	31.7	31.1	30.3	31.5	31.4
SD	50.3	48.8	42.8	54.3	49.6
Number of college courses in transition education					
< 3	69%	68	67	69	71
3-10	25	26	25	22	23
> 10	6	7	8	9	6
Number of professional development days in transition education					
< 10	66%	64	62	67	63
10-20	17	16	18	16	17
> 20	20	20	20	17	20
Ethnicity					
White or Caucasian	73%	74	73	72	72
Black or African-American	16	15	16	16	16
American Indian or Alaska Native	7	7	8	7	7
Mexican, Mexican-American, or Chicano	<1	< 1	< 1	< 1	< 1
Hispanic or Latino	3	2	3	3	3
Asian	1	1	1	1	<1

Table 44

Family Participants by Grade Focus Group

Characteristic	Grade Level			
	9th	10th	11th	12th
n	105	122	289	241
Age				
<18	2%	4	12	8
18-24	0	<1	2	3
25-29	<1	<1	1	<1
30-34	5	6	3	3
35-39	31	16	19	18
40-44	23	30	26	24
45-49	15	22	18	19
50-54	12	8	10	10
55-59	4	7	4	7
> 60	7	6	4	6
Mean	43.7	43.7	43.5	44.5
SD	8.4	10.6	9.1	9.5
Primary Language				
English	98	98	95	97
Other	2	2	5	3
Relationship to Student				
Father	14 %	19	11	10
Grandfather	0	2	1	<1
Brother	0	0	<1	<1
Male Guardian	0	0	1	<1
Other	0	0	2	2
Mother	76	75	76	77
Grandmother	7	4	4	4
Aunt	<1	0	<1	<1
Sister	0	0	1	<1
Female Guardian	<1	<1	1	3
Stepfather	<1	0	0	<1
Stepmother	0	0	2	0

Table 44

Family Participants by Grade Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th
Education				
Less than high school	7%	11	16	16
High school diploma/GED	33	49	42	39
Vocational certificate	18	11	13	9
Associate's degree	19	14	11	17
Bachelor's degree	16	13	13	15
Master's degree	6	3	3	4
Doctorate/Professional degree	2	0	1	<1
Employment				
Employed full-time	48	51	56	48
Employed part-time	11	5	8	10
Self-employed full-time	4	4	4	5
Self-employed part-time	4	2	2	3
Not working	25	30	22	22
Retired	9	8	5	6
Permanently disabled, not working for pay	0	0	2	7
Received help completing *TAGG*	3%	5	8	10
Help received				
Reading	0%	0	33	33
Writing	0	0	17	0
Translating	0	0	8	9
Explanation	100	0	25	55
Other	0	0	17	9
Ethnicity				
White	75%	72	63	60
Black or African-American	7	6	12	12
American Indian or Alaska Native	9%	11	8	11
Mexican, Mexican-American, Chicano	5%	<1	6	6
Puerto Rican	0	0	1	<1
Cuban	0	0	0	<1

Table 44

Family Participants by Grade Focus Group, continued

Characteristic	9th	10th	11th	12th
	Grade Level			
Hispanic or Latino	4	8	7	7
Native Hawaiian	0	0	0	0
Asian	0	0	<1	<1
Other	0	0	<1	2
Egyptian	0	0	0	0
German	1	0	0	0
Spanish	0	0	0	0
Middle Eastern	0	0	1	0
Multiethnic	0	<1	0	0

Table 45

Family Participants by Grade Non Focus Group

Characteristic		9th	10th	11th	12th
		Grade Level			
	n	688	675	508	556
Age					
<18		9%	9	6	8
18-24		2	2	2	1
25-29		1	1	1	1
30-34		3	3	4	4
35-39		18	21	20	20
40-44		25	24	25	25
45-49		20	18	18	19
50-54		10	12	11	11
55-59		3	6	7	5
> 60		5	6	6	6
	Mean	44.0	44.0	44.4	43.9
	SD	10.2	9.8	10.0	9.8
Primary Language					
English		96	96	98	97
Other		4	4	2	3

Table 45

Family Participants by Grade Non Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th
Relationship to Student				
Father	13%	15	14	14
Grandfather	1	1	< 1	1
Brother	< 1	< 1	< 1	< 1
Male Guardian	< 1	< 1	< 1	< 1
Other	1	2	1	< 1
Mother	76	75	76	76
Grandmother	4	4	5	4
Aunt	< 1	< 1	< 1	< 1
Sister	< 1	< 1	< 1	< 1
Female Guardian	2	2	2	1
Stepfather	< 1	< 1	< 1	< 1
Stepmother	< 1	< 1	< 1	1
Education				
Less than high school	15%	14	13	13
High school diploma/GED	43	39	41	42
Vocational certificate	11	12	12	14
Associate's degree	13	14	17	13
Bachelor's degree	14	14	14	14
Master's degree	3	4	4	4
Doctorate/Professional degree	1	1	< 1	1
Employment				
Employed full-time	52%	52	50	52
Employed part-time	8	9	8	8
Self-employed full-time	4	4	4	4
Self-employed part-time	2	3	3	3
Not working	23	22	25	3
Retired	6	6	7	7
Permanently disabled, not working for pay	4	4	4	1

Table 45

Family Participants by Grade Non Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th
Received help completing *TAGG*	7%	7	6	6
Help received				
Reading	35%	33	37	37
Writing	7	7	0	11
Translating	14	13	7	16
Explanation	35	39	60	26
Other	10	10	7	11
Ethnicity				
White	64%	64	67	68
Black or African-American	11	11	10	9
American Indian or Alaska Native	10	9	10	9
Mexican, Mexican-American, Chicano	5	6	4	4
Puerto Rican	<1	<1	<1	<1
Cuban	< 1	< 1	< 1	< 1
Hispanic or Latino	8	7	7	7
Native Hawaiian	0	< 1	0	< 1
Asian	2	2	2	1
Other	< 1	< 1	< 1	< 1
Egyptian	< 1	< 1	< 1	< 1
German	< 1	< 1	<1	< 1
Spanish	< 1	< 1	0	0
Middle Eastern	< 1	< 1	< 1	< 1
Multiethnic	< 1	0	< 1	< 1

Table 46

Student Participants by Grade Focus Group

Characteristic		Grade Level				
		9th	10th	11th	12th	18-21
	n	112	148	417	575	63
Age						
14		26%	0	0	0	0
15		50	21	0	0	0
16		14	50	17	0	0
17		4	23	51	13	3
18-21		4	7	31	87	97
	Mean	15.1	16.2	17.2	18.1	19.3
	SD	1.0	0.9	0.7	0.7	1.0
Gender						
Male		60%	52	57	61	68
Female		40	48	43	39	32
Primary Disability						
Autism		4%	2	5	2	24
Deaf-Blindness		0	0	< 1	0	0
Emotional Disturbance		10	9	5	3	11
Hearing Impaired		2	< 1	< 1	1	0
Intellectual Disability		7	9	12	12	28
Multiple Disability		0	< 1	< 1	2	5
Orthopedic Impairment		< 1	2	0	< 1	3
Other Health Impairment		19	9	16	13	8
Specific Learning Disability		52	63	58	62	12
Speech/Language Disability		2	2	2	2	0
Traumatic Brain Injury		2	0	< 1	< 1	3
Visual Impairment/Blindness		0	0	< 1	< 1	4
Other		2	1	< 1	< 1	1
Level of Disability						
Mild-Moderate		89%	81	91	92	85
Severe-Profound		2	4	3	5	15
Not Indicated		8	16	6	3	0
Attendance						
Absent $\leq$ 1 day/month		70%	65	59	56	66
Absent 2-3 days/month		20	25	58	32	31
Absent $\geq$ 4 days/month		10	11	13	12	3

Table 46

Student Participants by Grade Focus Group, continued

	Grade Level				
Characteristic	9th	10th	11th	12th	18-21
Attended IEP meeting	74%	90	86	91	93
Actively participated in IEP meeting	32	49	62	40	61
Led IEP meeting	2	4	10	12	19
Years teacher has known student					
< 1 year	83%	60	29	21	40
1-2 years	6	30	47	23	60
2-3 years	7	4	14	36	20
3-4 years	0	2	3	5	14
4-5 years	2	2	5	9	6
> 5 years	1	2	2	5	0
Mean	1.3	1.7	2.5	3.2	2.1
SD	1.7	1.6	2.0	2.5	1.4
Years student has attended school					
< 1 year	73%	14	12	11	36
1-2 years	9	60	9	4	20
2-3 years	4	14	52	11	7
3-4 years	< 1	5	14	50	7
4-5 years	0	2	4	4	7
> 5 years	13	8	8	18	24
Mean	1.6	2.2	3.0	4.0	2.7
SD	2.6	1.9	2.0	2.7	2.1
Years student has been enrolled in district					
< 1 year	22%	12	12	10	13
1-2 years	3	5	7	5	3
2-3 years	4	5	4	7	1
3-4 years	0	3	4	2	3
4-5 years	43	39	24	20	16
> 5 years	27	36	49	56	63
Mean	6.1	7.0	7.5	8.4	10.1
SD	4.2	4.0	4.3	4.6	4.7
Eligible for free/reduced lunch	52%	59	55	62	42
Receives free/reduced lunch	54	64	58	61	36
Receives ESL services	0	< 1	4	6	1

Table 46

Student Participants by Grade Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th	18-21
Ethnicity					
American	0%	0	0	0	1
American Indian or Alaska Native	10	7	8	12	0
Arabic	0	1	0	0	0
Asian	< 1		< 1	< 1	3
Black or African-American	11	14	24	22	10
Cuban	0	0	< 1	0	0
Egyptian	0	0	0	1	0
Guamanian, Chamorro				1	0
Hispanic or Latino	9	9	9	8	5
Irish	0	0		1	0
Multiethnic	0	0	< 1	< 1	0
Mexican, Mexican-American, or Chicano	3	5	5	8	15
Native Hawaiian	0	0	< 1	0	0
Other (Not Indicated)	0	0	0	0	0
Puerto Rican	0	0	< 1	0	0
White or Caucasian	66	63	51	47	65

Table 47

Student Participants by Grade Non Focus Group

Characteristic		Grade Level 9th	10th	11th	12th	18-21
	n	1239	1204	927	769	1289
Age						
14		0%	3	3	4	3
15		3	5	10	12	7
16		12	7	13	22	13
17		27	25	42	34	27
18-21		58	60	32	28	57
	Mean	17.6	17.6	17.5	16.8	17.3
	SD	1.1	1.2	1.5	1.3	1.2
Gender						
Male		58	59	60	57	59
Female		52	41	40	43	42

Table 47

Student Participants by Grade Non Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th	18-21
Primary Disability					
Autism	4%	4	4	6	3
Deaf-Blindness	< 1	< 1	< 1	< 1	< 1
Emotional Disturbance	5	5	6	7	5
Hearing Impaired	1	1	1	1	1
Intellectual Disability	12	12	12	11	12
Multiple Disability	1	1	1	1	1
Orthopedic Impairment	1	1	1	1	1
Other Health Impairment	< 1	< 1	< 1	< 1	< 1
Specific Learning Disability	14	15	13	14	14
Speech/Language Disability	60	58	58	56	61
Traumatic Brain Injury	2	1	1	2	2
Visual Impairment/Blindness	< 1	< 1	< 1	< 1	< 1
Other	< 1	< 1	< 1	< 1	< 1
Level of Disability					
Mild-Moderate	90	91	89	88	90
Severe-Profound	5	4	6	8	6
Not Indicated	4	5	5	4	4
Attendance					
Absent ≤ 1 day/month	58%	58	60	62	59
Absent 2-3 days/month	30	30	30	27	29
Absent ≥ 4 days/month	12	12	11	11	12
Attended IEP meeting	89%	88	89	85	88
Actively participated in IEP meeting	65	64	63	56	62
Led IEP meeting	10	10	10	8	9
Years teacher has known student					
< 1 year	26%	31	36	45	33
1-2 years	34	30	21	35	31
2-3 years	25	25	27	12	23
3-4 years	4	4	4	3	4
4-5 years	7	7	7	4	6
> 5 years	4	3	4	2	3
Mean	1.6	1.5	1.6	1.0	1.5
SD	1.9	1.9	2.0	1.5	1.9

Table 47

Student Participants by Grade Non Focus Group, continued

Characteristic	Grade Level 9th	10th	11th	12th	18-21
Years student has attended school					
< 1 year	13%	20	23	26	18
1-2 years	15	8	17	21	14
2-3 years	26	25	11	32	25
3-4 years	29	29	31	9	27
4-5 years	4	4	4	3	3
> 5 years	13	14	15	10	13
Mean	2.4	2.3	2.3	1.8	2.3
SD	1.5	1.6	1.7	1.5	1.5
Years student has been enrolled in district					
< 1 year	8%	9	9	11	9
1-2 years	6	6	5	6	6
2-3 years	6	6	7	5	6
3-4 years	3	3	2	3	2
4-5 years	25	25	27	30	27
> 5 years	53	53	52	46	50
Mean	6.5	6.4	6.4	6.0	6.2
SD	3.4	3.9	3.8	3.8	3.8
Eligible for free/reduced lunch	58%	58	59	55	58
Receives free/reduced lunch	58	58	58	57	59
Receives ESL services	4	4	4	2	4

Table 47

Student Participants by Grade Non Focus Group, continued

Characteristic	Grade Level				
	9th	10th	11th	12th	18-21
Ethnicity					
American	< 1%	< 1	< 1	< 1	< 1
American Indian or Alaska Native	10	10	11	8	11
Arabic	< 1	< 1	< 1	< 1	< 1
Asian	2	2	2	2	2
Black or African-American	21	21	18	20	20
Cuban	< 1	< 1	< 1	< 1	< 1
Egyptian	< 1	< 1	< 1	< 1	< 1
Guamanian, Chamorro	< 1	< 1	< 1	< 1	< 1
Hispanic or Latino	6	7	7	9	8
Irish	< 1	< 1	< 1	0	< 1
Multiethnic	< 1	< 1	< 1	< 1	< 1
Mexican, Mexican-American, or Chicano	6	6	8	6	6
Native Hawaiian	< 1	< 1	< 1	< 1	< 1
Other (Not Indicated)	0	0	0	0	< 1
Puerto Rican	< 1	< 1	< 1	< 1	< 1
White or Caucasian	51	51	54	56	52

Data analysis of grade level groups. This analysis followed the steps described for the standard deviation groups and examined the properties of the *TAGG* professional and student versions across grade levels. To begin, the factor structure for all three versions of the *TAGG* was applied to its respective groups to understand whether the same items were associated with the same constructs for each group. Then, each group was assessed for whether the strengths of the item-constructs relations were consistent with those of the full *TAGG* sample. Again, the Family version was not included because when all three annual phases of family data were combined, the eight-factor structure did not appear to be suitable.

Disability Category Labels

Disability categorical labels are used primarily to classify students and to designate some services. Students are evaluated every three years, and may change categories depending on academic progress and IEP team members. However, a student's *TAGG* score should not change as a function of his disability category, as was expected to occur for the other variables considered for this study.

Participants. Out of obligation to traditional special education practices, the data will be analyzed across the following disability groups: Autism, ED, ID, OHI, and SLD as the data set contains enough participants from each of these groups to make the analysis possible. Tables 48-53 include the demographic information for participants in these categories. Worth noting is that the students were all about the same age, and about 90% of the students in each of the categories were described as having mild/moderate level by the professional who completed the *TAGG* form on the student.

Table 48

Educator Participants by Disability Category Focus Group

Characteristic		Autism	ED	ID	OHI	SLD
		Disability Group				
	n	61	79	174	200	825
Age						
18-24		3%	4	4	3	4
25-29		5	16	3	6	5
30-34		5	7	7	8	8
35-39		8	10	7	11	10
40-44		16	19	13	23	15
45-49		3	9	9	5	8
50-54		22	15	19	15	16
55-59		24	12	24	22	26
> 60		14	3	13	6	8
	Mean	44.5	42.0	47.5	43.6	45.9
	SD	11.0	10.1	10.4	10.1	10.4
Gender						
Male		8%	6	4	6	4
Female		92	94	96	96	94
Level of Education						
Bachelor's Degree		5%	6	10	14	9
Some Master's Coursework		26	17	32	26	25
Master's Degree		57	59	43	49	55
Educational Specialist		5	6	3	7	8
Some Doctoral Coursework		7	12	10	5	3
Ph.D. or Ed.D.		0	0	2	< 1	< 1
Years teaching students with disabilities						
< 3		5%	13	5	10	9
3-10		38	46	34	35	33
> 10		57	42	61	55	58
	Mean	13.5	12.4	15.8	14.2	15.3
	SD	9.2	10.1	9.5	10.3	10.6

Table 48

Educator Participants by Disability Category Focus Group, continued

Characteristic		Disability Group			
	Autism	ED	ID	OHI	SLD
Years teaching at current school					
< 3	13%	34	18	27	24
3-10	57	44	61	55	47
> 10	30	22	21	19	30
Mean	8.0	7.4	7.0	7.4	8.9
SD	5.9	7.3	4.4	6.8	7.5
Years teaching in current district					
< 3	36%	56	45	56	56
3-10	33	25	29	35	28
> 10	31	20	23	10	16
Mean	10.4	9.2	11.1	9.8	11.4
SD	7.2	8.6	7.8	8.6	9.4
Lives in district	53%	54	56	60	58
Number of students on caseload					
< 5	10%	10	10	5	7
5-10	12	4	9	6	5
10-15	23	23	14	18	17
> 15	56	63	66	73	71
Mean	36.7	26.4	252	29.0	32.5
SD	108.5	39.6	28.9	42.0	46.0
Number of college courses in transition education					
< 3	65%	56	55	75	73
3-10	28	25	39	20	21
> 10	7	20	6	5	6

Table 48

Educator Participants by Disability Category Focus Group, continued

Characteristic	Disability Group				
	Autism	ED	ID	OHI	SLD
Number of professional development days in transition education					
< 10	69%	57	62	66	65
10-20	16	28	16	15	16
> 20	15	15	22	20	20
Ethnicity					
White or Caucasian	90%	55	79	75	71
Black or African-American	5	25	15	12	16
American Indian or Alaska Native	2	8	2	11	8
Mexican, Mexican-American, or Chicano	2	1	< 1	1	4
Hispanic or Latino	0	0	0	0	< 1
Asian	2	0	0	2	< 1

Table 49

Educator Participants by Disability Category Non Focus Group

Characteristic		Disability Group				
		Autism	ED	ID	OHI	SLD
	n	1391	1373	1278	1252	624
Age						
18-24		4%	4	4	4	4
25-29		5	5	6	5	5
30-34		9	9	9	9	9
35-39		9	9	9	9	8
40-44		16	16	17	15	17
45-49		8	8	8	8	7
50-54		17	17	17	18	16
55-59		24	25	23	24	24
> 60		8	9	8	9	8
	Mean	45.5	45.6	45.1	45.7	44.8
	SD	10.4	10.4	10.4	10.5	10.4

Table 49

Educator Participants by Disability Category Non Focus Group, continued

Characteristic		Disability Group			
	Autism	ED	ID	OHI	SLD
Gender					
Male	5%	5	5	5	4
Female	95	95	95	95	96
Level of Education					
Bachelor's Degree	10%	10	9	9	10
Some Master's Coursework	25	25	24	25	26
Master's Degree	54	53	55	54	54
Educational Specialist	7	7	7	7	6
Some Doctoral Coursework	5	5	4	5	5
Ph.D. or Ed.D.	<1	<1	<1	<1	<1
Years teaching students with disabilities					
< 3	8%	8	9	8	8
3-10	35	35	35	35	34
> 10	57	58	56	57	57
Mean	15.1	15.2	14.9	15.1	15.0
SD	10.4	10.4	10.5	10.4	10.5
Years teaching at current school					
< 3	24%	23	25	23	23
3-10	50	51	49	50	52
> 10	26	26	27	27	26
Mean	8.3	8.3	8.4	8.4	8.3
SD	7.1	7.0	7.3	7.1	7.1
Years teaching in current district					
< 3	53%	52	54	52	51
3-10	30	30	30	29	30
> 10	17	18	17	19	19
Mean	10.9	11.0	10.9	11.1	11.0
SD	9.1	9.0	9.2	9.1	9.0

Table 49

Educator Participants by Disability Category Non Focus Group, continued

Characteristic		Disability Group			
	Autism	ED	ID	OHI	SLD
Lives in district	58%	58	58	58	59
Number of students on caseload					
< 5	8%	7	7	8	7
5-10	5	6	5	6	6
10-15	17	17	18	17	18
> 15	70	70	70	69	70
Mean	31.0	31.5	32.1	31.6	31.7
SD	44.8	49.7	51.4	50.3	50.1
Number of college courses in transition education					
< 3	69%	69	71	68	69
3-10	24	24	22	25	25
> 10	7	6	7	7	6
Number of professional development days in transition education					
< 10	63%	64	64	63	63
10-20	17	16	17	17	17
> 20	20	20	19	19	20
Ethnicity					
White or Caucasian	72%	73	72	73	73
Black or African-American	17	15	16	17	16
American Indian or Alaska Native	7	7	8	7	7
Mexican, Mexican-American, or Chicano	<1	<1	<1	<1	<1
Hispanic or Latino	3	3	3	3	2
Asian	<1	1	1	<1	<1

Table 50

Family Participants by Disability Category Focus Group

Characteristic		Disability Group		
		ID	OHI	SLD
	n	97	109	442
Age				
<18		9%	9	9
18-24		3	2	2
25-29		0	0	1
30-34		3	4	3
35-39		15	18	21
40-44		20	25	27
45-49		24	22	19
50-54		11	10	9
55-59		9	9	4
> 60		5	<1	5
	Mean	44.5	44.1	43.7
	SD	10.0	8.1	8.8
Primary Language				
English		96%	99	96
Other		4	<1	4
Relationship to Student				
Father		14%	12	11
Grandfather		0	0	1
Brother		0	0	<1
Male Guardian		1	0	<1
Other		2	2	1
Mother		73	82	78
Grandmother		6	2	4
Aunt		0	0	<1
Sister		0	0	<1
Female Guardian		2	0	1
Stepfather		0	<1	<1
Stepmother		1	<1	<1

Table 50

Family Participants by Disability Category Focus Group, continued

Characteristic	Disability Group ID	OHI	SLD
Education			
Less than high school	16%	9	16
High school diploma/GED	37	40	43
Vocational certificate	9	12	12
Associate's degree	18	15	15
Bachelor's degree	14	19	10
Master's degree	3	3	4
Doctorate/Professional degree	3	2	<1
Employment			
Employed full-time	38%	61	54
Employed part-time	7	8	8
Self-employed full-time	2	3	4
Self-employed part-time	0	6	3
Not working	42	15	22
Retired	6	3	7
Permanently disabled, not working for pay	5	3	2
Received help completing *TAGG*	15%	6	7
Help received			
Reading	33%	25	43
Writing	33	0	5
Translating	0	25	14
Explanation	33	50	29
Other	0	0	10
Ethnicity			
White	55%	79	63
Black or African-American	20	2	11
American Indian or Alaska Native	10	10	10
Mexican, Mexican-American, Chicano	1	4	6
Puerto Rican	0	1	<1
Cuban	1	0	0
Hispanic or Latino	6	2	9
Native Hawaiian	0	0	0

Table 50

Family Participants by Disability Category Focus Group, continued

Characteristic	Disability Group ID	OHI	SLD
Asian	<1%	0	<1
Other	0	0	<1
Egyptian	<1	0	0
German	0	0	<1
Spanish	0	0	0
Middle Eastern	0	0	<1
Multiethnic	<1	0	0

Table 51

Family Participants by Disability Category Non Focus Group

Characteristic	Disability Group ID	OHI	SLD
n	696	684	351
Age			
<18	8%	8	7
18-24	2	2	2
25-29	1	1	1
30-34	4	4	4
35-39	21	20	20
40-44	25	25	25
45-49	18	18	18
50-54	11	11	12
55-59	5	6	5
> 60	6	6	7
Mean	43.9	44.1	44.6
SD	10.0	9.9	10.4
Primary Language			
English	97%	95	97
Other	3	4	3

Table 51

Family Participants by Disability Category Non Focus Group, continued

Characteristic	Disability Group ID	OHI	SLD
Relationship to Student			
Father	13%	13	14
Grandfather	1	1	< 1
Brother	< 1	< 1	0
Male Guardian	< 1	< 1	0
Other	1	1	1
Mother	75	74	75
Grandmother	4	5	5
Aunt	< 1	< 1	< 1
Sister	< 1	< 1	< 1
Female Guardian	2	2	3
Stepfather	< 1	< 1	< 1
Stepmother	< 1	< 1	< 1
Education			
Less than high school	13%	14	11
High school diploma/GED	41	40	37
Vocational certificate	12	12	12
Associate's degree	13	14	13
Bachelor's degree	13	13	20
Master's degree	4	4	4
Doctorate/Professional degree	< 1	< 1	2
Employment			
Employed full-time	54%	51	48
Employed part-time	8	8	10
Self-employed full-time	4	4	5
Self-employed part-time	3	2	3
Not working	21	25	25
Retired	6	7	5
Permanently disabled, not working for pay	3	3	4

Table 51

Family Participants by Disability Category Non Focus Group, continued

Characteristic	Disability Group ID	OHI	SLD
Received help completing *TAGG*	6%	7	7
Help received			
Reading	36%	37	23
Writing	4	10	15
Translating	14	10	8
Explanation	36	33	46
Other	10	10	8
Ethnicity			
White	67%	64	69
Black or African-American	8	8	9
American Indian or Alaska Native	8	10	9
Mexican, Mexican-American, Chicano	5	5	3
Puerto Rican	< 1	<1	< 1
Cuban	6	7	6
Hispanic or Latino	< 1	< 1	< 1
Native Hawaiian	0	0	0
Asian	< 1	2	< 1
Other	0	< 1	0
Egyptian	0	< 1	< 1
German	0	< 1	0
Spanish	0	0	< 1
Middle Eastern	0	< 1	0
Multiethnic	0	< 1	0

Table 52

Student Participants by Disability Category Focus Group

Characteristic		Disability Group				
		Autism	ED	ID	OHI	SLD
	n	55	71	154	189	778
Age						
14		2%	9	1	2	1
15		10	13	4	9	6
16		3	21	10	9	14
17		32	11	20	31	26
18-21		53	44	63	48	52
	Mean	18	17	18	17	17
	SD	1.6	1.8	1.25	1.3	1.2
Gender						
Male		82%	64	47	61	59
Female		18	46	53	39	41
Level of Disability						
Mild-Moderate		87%	92	94	90	90
Severe-Profound		10	5	4	5	2
Not Indicated		3	3	2	4	7
Attendance						
Absent ≤ 1 day/month		90%	46	60	67	57
Absent 2-3 days/month		7	37	29	24	31
Absent ≥ 4 days/month		3	18	10	12	12
Attended IEP meeting		90%	86	88	81	90
Actively participated in IEP meeting		50	58	54	64	66
Led IEP meeting		12	12	11	8	9
Years teacher has known student						
< 1 year		26%	54	31	40	30
1-2 years		23	26	21	29	33
2-3 years		29	14	27	20	24
3-4 years		14	0	6	6	2
4-5 years		9	4	10	4	6
>5 years		0	2	4	2	4
	Mean	2.7	2.0	3.1	2.5	2.6
	SD	1.5	2.5	2.8	2.1	2.4

Table 52

Student Participants by Disability Category Focus Group, continued

Characteristic	Disability Group				
	Autism	ED	ID	OHI	SLD
Years student has attended school					
< 1 year	18%	40	25	18	17
1-2 years	14	18	11	14	13
2-3 years	14	23	12	21	27
3-4 years	18	6	28	29	27
4-5 years	7	2	7	5	3
>5 years	30	11	17	11	13
Mean	3.7	2.2	3.3	3.2	3.3
SD	2.8	2.0	2.4	2.5	2.7
Years student has been enrolled in district					
< 1 year	9%	24	18	8	11
1-2 years	6	9	5	4	6
2-3 years	4	2	4	5	6
3-4 years	8	3	<1	2	3
4-5 years	28	24	18	25	26
>5 years	45	38	54	57	48
Mean	8.3	6.6	8.4	8.6	7.6
SD	4.2	4.8	4.9	4.4	4.4
Eligible for free/reduced lunch	31%	59	72	51	59
Receives free/reduced lunch	32	61	74	50	59
Receives ESL services	3	0	4	<1	4

Table 52

Student Participants by Disability Category Focus Group, continued

Characteristic	Disability Group				
	Autism	ED	ID	OHI	SLD
Ethnicity					
American	0	0	0	0	<1
American Indian or Alaska Native	4	6	7	12	11
Arabic	0	0	0	<1	0
Asian	0	0	1	<1	0
Black or African-American	7	22	33	14	19
Cuban	0	0	0	1	0
Egyptian	0	0	<1	0	0
Guamanian, Chamorro	0	0	<1	0	0
Hispanic or Latino	11	6	3	6	11
Irish	0	0	0	0	0
Multiethnic	0	0	0	<1	<1
Mexican, Mexican-American, or Chicano	9	3	5	4	7
Native Hawaiian	0	0	0	<1	0
Other (Not Indicated)	0	1	<1	2	<1
Puerto Rican	0	2	0	<1	<1
White or Caucasian	70	59	50	60	50

Table 53

Student Participants by Disability Category Non Focus Group

Characteristic	Disability Group				
	Autism	ED	ID	OHI	SLD
n	1301	1285	1202	1167	578
Age					
14	2%	2	2	2	2
15	7	7	7	6	8
16	13	12	13	13	11
17	25	26	26	24	25
18-21	53	53	53	55	55
Mean	17.4	17.4	17.4	17.5	17.5
SD	1.2	1.3	1.3	1.3	1.4

Table 53

Student Participants by Disability Category Non Focus Group, continued

	Disability Group				
Characteristic	Autism	ED	ID	OHI	SLD
Gender					
Male	58%	58	58	60	59
Female	42	42	42	40	41
Primary Disability					
Autism	0%	5	5	5	10
Deaf-Blindness	< 1	< 1	< 1	< 1	< 1
Emotional Disturbance	6	0	6	6	13
Hearing Impaired	1	1	1	1	2
Intellectual Disability	12	12	0	12	24
Multiple Disability	1	1	1	1	2
Orthopedic Impairment	1	1	1	1	2
Speech/Language Disability	< 1	< 1	< 1	< 1	1
Other Health Impairment	16	16	16	0	37
Specific Learning Disability	67	67	67	61	0
Traumatic Brain Injury	2	2	2	2	4
Visual Impairment/Blindness	< 1	< 1	< 1	< 1	2
Other	< 1	< 1	< 1	< 1	1
Level of Disability					
Mild-Moderate	90%	90	89	90	89
Severe-Profound	5	5	6	6	4
Not Indicated	6	6	5	4	7
Attendance					
Absent ≤ 1 day/month	58%	60	58	59	63
Absent 2-3 days/month	30	29	30	30	26
Absent ≥ 4 days/month	12	11	12	11	11
Attended IEP meeting	88%	88	88	89	89
Actively participated in IEP meeting	63	63	63	62	58
Led IEP meeting	9	9	9	10	10

Table 53

Student Participants by Disability Category Non Focus Group, continued

Characteristic	Autism	ED	ID	OHI	SLD
	Disability Group				
Years teacher has known student					
< 1 year	34%	32	34	33	37
1-2 years	30	30	33	31	26
2-3 years	23	24	22	24	21
3-4 years	4	4	4	4	5
4-5 years	6	6	6	7	7
>5 years	4	3	3	4	3
Mean	1.5	1.5	1.4	1.4	1.4
SD	2.0	2.0	1.9	1.8	1.9
Years student has attended school					
< 1 year	20%	18	19	20	22
1-2 years	14	14	15	14	15
2-3 years	24	24	25	24	20
3-4 years	26	28	25	25	24
4-5 years	4	4	3	4	5
>5 years	12	13	13	13	14
Mean	2.2	2.2	2.2	2.3	2.2
SD	1.6	1.6	1.6	1.5	1.6
Years student has been enrolled in district					
< 1 year	9%	8	9	10	9
1-2 years	6	5	6	6	5
2-3 years	6	6	6	6	6
3-4 years	2	2	3	3	2
4-5 years	26	26	27	26	25
>5 years	51	51	50	50	53
Mean	6.3	6.3	6.3	6.5	6.6
SD	3.9	3.8	3.8	3.8	3.8
Eligible for free/reduced lunch	59%	58	56	59	57
Receives free/reduced lunch	59	58	56	59	57
Receives ESL services	4	4	4	4	2

Table 53

Student Participants by Disability Category Non Focus Group, continued

Characteristic	Disability Group				
	Autism	ED	ID	OHI	SLD
Ethnicity					
American	< 1	< 1	< 1	< 1	0
American Indian or Alaska Native	11	10	11	10	9
Arabic	< 1	< 1	< 1	< 1	< 1
Asian	2	1	2	2	3
Black or African-American	20	20	18	21	20
Cuban	< 1	< 1	< 1	< 1	< 1
Egyptian	< 1	< 1	< 1	< 1	< 1
Guamanian, Chamorro					
Hispanic or Latino	8	7	7	7	6
Irish	< 1	< 1	< 1	< 1	0
Multiethnic	< 1	< 1	< 1	< 1	< 1
Mexican, Mexican-American, or Chicano	10	9	9	< 1	< 1
Native Hawaiian	< 1	< 1	< 1	< 1	< 1
Other (Not Indicated)	0	0	0	0	< 1
Puerto Rican	< 1	< 1	< 1	< 1	< 1
White or Caucasian	52	52	53	51	58

Data analysis of disability category groups. This analysis followed the procedures described in the standard deviation group for answering research questions one and two across disability categories. The first step in the analysis entailed applying the factor structure for each version of the *TAGG* to each of the disability groups to understand whether the groups experienced the same item-construct relations. If the structure was suitable for each group, the strength of the item-construct relations was examined to ensure that *TAGG* scores were a function of the skills it intends to assess, not disability category.

CHAPTER FOUR

Results

The degree to which students with disabilities participate in general education is impacted by a variety of factors including their level of need/degree of disability, the model of inclusion used by their schools, and the degree to which their families advocate for their placement. Regardless of the reasons for their placement, students' access to special education and general education services should be considered on any assessment that seeks to understand behaviors that are more likely to be learned in one setting than in another—such as disability awareness and IEP involvement.

This study was designed to understand whether participating in general education contributed to students' scores on the *Transition Assessment and Goal Generator*, and whether the percent of time spent in general education needed to be considered when goals were suggested for inclusion on the *TAGG* scoring profile. To do this, I examined each version of the *TAGG* to understand whether the assessment items were paired with the same constructs for students who had various levels of access to the *TAGG* skills, and then examined whether the items had the same strength of relations across levels of access.

Data Normality and Missing Data

Prior to the multigroup analysis, the normality of the *TAGG* data set was assessed to ensure the continuous items met the assumptions of CFA. Three items on the *TAGG* are categorical (GA6, E3, and E4), and therefore did not need to be tested for meeting this assumption. Item means, standard deviations, skewness and kurtosis for the 31 continuous items are presented in Tables 54-56.

Table 54

Item Means, Standard Deviations, Skewness, and Kurtosis for TAGG-P

Item	Mean	SD	Skewness	Kurtosis
SL1	2.73	1.06	-0.68	-0.04
SL2	2.53	1.10	-0.46	-0.45
SL3	2.40	1.08	-0.29	-0.58
SL4	2.70	1.05	-0.56	-0.27
DA1	2.42	1.25	-0.46	-0.71
DA2	2.29	1.15	-0.26	-0.67
DA3	2.71	1.21	-0.73	-0.34
DA4	1.93	1.35	0.01	-1.17
P1	2.94	1.10	-0.86	-0.06
P2	2.67	1.12	-0.48	-0.61
P3	2.37	1.14	-0.21	-0.79
P4	2.40	1.18	-0.28	-0.81
P5	2.55	1.10	-0.38	-0.56
IO1	2.84	1.10	-0.73	-0.25
IO2	1.99	1.52	0.02	-1.44
IO3	3.06	1.03	-0.89	0.00
GA1	2.44	1.10	-0.38	-0.53
GA2	2.10	1.11	-0.08	-0.68
GA3	2.06	1.14	-0.09	-0.76
GA4	1.94	1.13	-0.01	-0.74
GA5	2.24	1.12	-0.24	-0.68
E1	3.11	1.11	-1.17	0.55
E2	2.85	1.16	-0.79	-0.22
IEP1	2.63	1.34	-0.68	-0.71
IEP2	1.76	1.34	0.11	-1.18
IEP3	1.58	1.36	0.32	-1.16
IEP4	0.78	1.13	1.36	0.86
CS1	2.62	1.13	-0.51	-0.50
CS2	3.05	0.95	-0.83	0.13
CS3	2.66	1.16	-0.53	-0.59
CS4	1.50	1.32	0.41	-1.01

Table 55

Item Means, Standard Deviations, Skewness, and Kurtosis for TAGG-F

Item	Mean	SD	Skewness	Kurtosis
SL1	2.49	1.17	-0.43	-0.54
SL2	2.24	1.21	-0.17	-0.88
SL3	2.51	1.22	-0.42	-0.79
SL4	2.43	1.17	-0.39	-0.64
DA1	1.48	1.28	0.43	-0.87
DA2	1.53	1.32	0.40	-1.01
DA3	2.13	1.38	-0.19	-1.17
DA4	1.53	1.40	0.40	-1.15
P1	3.07	1.19	-1.15	0.33
P2	2.61	1.17	-0.41	-0.65
P3	2.37	1.21	-0.20	-0.87
P4	2.37	1.21	-0.22	-0.84
P5	2.63	1.13	-0.42	-0.58
IO1	2.92	1.16	-0.88	-0.03
IO2	2.14	1.56	-0.15	-1.49
IO3	3.29	1.00	-1.38	1.27
GA1	2.27	1.25	-0.28	-0.82
GA2	1.80	1.20	0.12	-0.71
GA3	2.04	1.22	0.00	-0.84
GA4	1.98	1.18	-0.01	-0.74
GA5	2.30	1.21	-0.23	-0.81
E1	3.30	1.08	-1.56	1.72
E2	3.06	1.17	-1.12	0.32
IEP1	2.85	1.29	-0.92	-0.21
IEP2	2.37	1.37	-0.39	-1.03
IEP3	2.15	1.44	-0.18	-1.26
IEP4	1.13	1.31	0.86	-0.47
CS1	2.90	1.25	-0.90	-0.24
CS2	2.97	1.05	-0.85	0.16
CS3	2.77	1.20	-0.71	-0.34
CS4	1.26	1.39	0.69	-0.86

Table 56

Item Means, Standard Deviations, Skewness, and Kurtosis for TAGG-S

Item	Mean	SD	Skewness	Kurtosis
CS_SL1	1.56	0.60	-1.01	0.02
CS_SL2	1.50	0.62	-0.86	-0.28
CS_SL3	1.24	0.72	-0.39	-1.00
CS_SL4	1.21	0.67	-0.27	-0.82
CS_SL5	1.56	0.64	-1.14	0.16
CS_SL6	1.45	0.63	-0.70	-0.50
CS_SL7	1.39	0.69	-0.68	-0.69
CS_SL8	0.76	0.81	0.46	-1.33
DA1	1.01	0.75	-0.01	-1.23
DA2	0.89	0.75	0.19	-1.19
DA3	1.09	0.79	-0.17	-1.39
DA4	0.85	0.80	0.28	-1.39
P1	1.66	0.61	-1.63	1.45
P2	1.51	0.63	-0.92	-0.20
P3	1.39	0.66	-0.62	-0.64
P4	1.38	0.65	-0.57	-0.67
P5	1.57	0.61	-1.11	0.17
IO1	1.42	0.68	-0.74	-0.59
IO2	1.00	0.84	0.01	-1.60
IO3	1.57	0.61	-1.13	0.20
GA1	1.29	0.68	-0.43	-0.83
GA2	1.14	0.67	-0.18	-0.81
GA3	1.29	0.66	-0.39	-0.77
GA4	1.24	0.69	-0.35	-0.89
GA5	1.48	0.66	-0.88	-0.35
E1	1.72	0.58	-1.94	2.59
E2	1.77	0.54	-2.28	4.10
IEP1	1.36	0.77	-0.72	-0.96
IEP2	1.22	0.79	-0.42	-1.27
IEP3	1.11	0.80	-0.21	-1.41
IEP4	0.55	0.74	0.93	-0.57

All of the items on all three versions of the *TAGG* were within the limits for normality as described by Kline (2011), who suggests skewness falls within -3.00 and 3.00, and kurtosis falls within -10.00 and 10.00.

Missing data was handled by imputing each student's construct mean when appropriate. Prior to making these imputations, the *TAGG-P* was missing about 1.12% of the responses, the *TAGG-F* was missing about 1.5% of the items, and the *TAGG-S* missing about 4.28% of the items. After imputing the means, the *TAGG-P* was missing about 0.84% of the items, the *TAGG-F* was missing 1.26% of the items, and the *TAGG-S* was missing 3.91% of the items. Tables 57-59 include the number of missing cases for each item before and after the imputations, along with the percent of imputations made for each item. A detailed description of this process can be found in Chapter 3.

Table 57

Percent Missing Before and After Mean Imputations TAGG-P

Item	Missing Before Imputation	Missing After Imputation	Percent Imputed
SL1	7	4	42.9
SL2	8	5	37.5
SL3	10	7	30.0
SL4	10	7	30.0
DA1	10	7	30.0
DA2	10	7	30.0
DA3	16	11	31.3
DA4	14	9	35.7
P1	15	10	33.3
P2	14	9	35.7
P3	14	9	35.7
P4	14	9	35.7
P5	20	15	25.0
IO1	16	11	31.3
IO2	15	10	33.3
IO3	14	9	35.7

Table 57

Percent Missing Before and After Mean Imputations TAGG-P, continued

Item	Missing Before Imputation	Missing After Imputation	Percent Imputed
GA1	16	13	18.8
GA2	12	9	25.0
GA3	13	10	23.1
GA4	12	9	25.0
GA5	15	12	20.0
GA6[a]	23	23	0.0
E1[b]	14	14	0.0
E2[b]	15	15	0.0
E3[a]	21	21	0.0
E4[a]	21	21	0.0
IEP1	22	16	27.3
IEP2	22	16	27.3
IEP3	22	16	27.3
IEP4	24	18	25.0
CS1	20	14	30.0
CS2	19	13	31.6
CS3	19	13	31.6
CS4	21	15	28.6
Total	538	407	24.3
Percent	1.1	0.8	

Note. [a]Items are binary/categorical and could not be imputed. [b]Construct did not contain adequate items for imputation.

Table 58

Percent Missing Before and After Mean Imputations TAGG-F

Item	Missing Before Imputation	Missing After Imputation	Percent Imputed
SL1	2	0	100.0
SL2	6	3	50.0
SL3	7	4	42.9
SL4	8	7	12.5
DA1	4	4	0.0
DA2	4	4	0.0

Table 58

Percent Missing Before and After Mean Imputations TAGG-F, continued

Item	Missing Before Imputation	Missing After Imputation	Percent Imputed
DA3	14	10	28.6
DA4	3	3	0.0
P1	6	2	66.7
P2	5	3	40.0
P3	5	3	40.0
P4	9	7	22.2
P5	8	5	37.5
IO1	9	6	33.3
IO2	5	3	40.0
IO3	4	2	50.0
GA1	4	2	50.0
GA2	5	5	0.0
GA3	9	7	22.2
GA4	7	5	28.6
GA5	9	7	22.2
GA6[a]	132	132	0.0
E1[b]	7	7	0.0
E2[b]	8	8	0.0
E3[a]	9	9	0.0
E4[a]	10	10	0.0
IEP1	18	13	27.8
IEP2	20	14	30.0
IEP3	15	13	13.3
IEP4	16	14	12.5
CS1	10	7	30.0
CS2	7	6	14.3
CS3	9	7	22.2
CS4	11	8	27.3
Total	405	341	15.8
Percent	1.50	1.26	

Note. [a]Items are binary/categorical and could not be imputed. [b]Construct did not contain adequate items for imputation.

Table 59

Percent Missing Before and After Mean Imputations TAGG-S

Item	Missing Before Imputation	Missing After Imputation	Percent Imputed
SC/SL1	44	39	11.4
SC/SL2	48	42	12.5
SC/SL3	52	45	13.5
SC/SL4	42	38	9.5
SC/SL5	50	43	14.0
SC/SL6	56	48	14.3
SC/SL7	52	45	13.5
SC/SL8	48	42	12.5
DA1	40	36	10.0
DA2	43	38	11.6
DA3	45	38	15.6
DA4	44	39	11.4
P1	40	35	12.5
P2	47	43	8.5
P3	40	35	12.5
P4	41	35	14.6
P5	44	40	9.1
IO1	44	40	9.1
IO2	40	35	12.5
IO3	39	35	10.3
GA1	49	42	14.3
GA2	41	36	12.2
GA3	52	47	9.6
GA4	49	43	12.2
GA5	66	58	12.1
GA6[a]	387	387	0.0
E1[b]	59	59	0.0
E2[b]	52	52	0.0
E3[a]	52	52	0.0
E4[a]	49	49	0.0
IEP1	58	48	17.2

Table 59

Percent Missing Before and After Mean Imputations TAGG-S, continued

Item	Missing Before Imputation	Missing After Imputation	Percent Imputed
IEP2	51	44	13.7
IEP3	55	48	12.7
IEP4	53	46	13.2
Total	1972	1802	8.6
Percent	4.3	3.9	

Note. [a]Items are binary/categorical and could not be imputed. [b]Construct did not contain adequate items for imputation.

One item from the *TAGG-F* and the *TAGG-S* was missing at a far greater proportion than other items. This item asked for a yes or no response to whether the student had attained at least one annual transition goal. No response was given to this item on 48.8% of the *TAGG-F* forms and 9.7% of *TAGG-S* forms, possibly because the participants did not understand the question or simply did not know the answer. Because the purpose of this study was to test the *TAGG* assessments as they stand, the item was included in the analysis. For this item and any other items that could not be imputed, pairwise deletion was used to retain as many cases as possible when the correlation matrix was computed. Other data appeared to be missing completely at random.

***TAGG* Professional Version**

I wanted to understand whether the item-construct relations on the *TAGG-P* were suitable across groups of students who spent varying degrees of time in general education. The *TAGG-P* has shown acceptable fit in development (χ^2 = 1043.62, *df* = 499; RMSEA = .058, CFI = .92, TLI = .91, and RMSR = .0597) and in replication (χ^2 =

2863.49, *df* = 1021, RMSEA = .072, CFI = .88, TLI = .88, RMSR = .065), though some differences have been observed in the strengths of the item-construct relations (Hennessey et al., 2013), potentially due to differences in the sampling procedures across versions. The results presented below reflect the inclusion of all phases of the *TAGG-P* collected during the first three phases of development in the analysis.

Multigroup analysis. I used a two-step process to understand whether the eight-factor *TAGG* configuration was appropriate for several groups of *TAGG* users. In the first step of the analysis, which answers research question one, I estimated the fit of the eight-construct model for the focus group (for example, professional educators using the *TAGG-P* to rate students who spend less than 25% of their day in general education) and the remaining participants separately and allowed the focus groups to have different factor loadings from the non-focus group. As part of this analysis, the first item for each construct (i.e., the conceptually easiest item) was constrained to one to scale the constructs.

For the second step of the analysis, which answered research question two, I constrained the factor loadings to be the same for both the focus and non-focus groups, and subsequently tested for a significant change in the overall fit. When the change statistics indicated a negative change, it could be understood that the suitability of the model decreased for the focus group (i.e. students who spent about 30-40% of their school day in general education) when the factor loadings were held equal between the groups. A significant decrease indicates the eight-factor structure is suitable for the focus group, but that the strengths of the item-factor relations are not the same for the focus group as for the other participants who used the *TAGG*.

Percent of time in general education groups. I divided into groups based on the percent of the day they spent in general education using the mean and half standard deviation intervals. Most of the students in this study spent more than half of their day in general education, which may have skewed the model in their favor during development. This portion of the study allows for a targeted examination of students who spend less time in general education.

Item-construct relations. This study found consistently good fit of the eight-factor *TAGG-P* model across participation in general education, meaning educators associated the same behaviors with the same constructs for students regardless of their participation in general education. For example, no matter how much time their students spend in general education, educators associated students using plans they made to attain their goals with the Goal Attainment construct.

The RMSEAs for the groups ranged from .061 for students who spent 25-29% of their day in general education to .057 for students who spent 39.3-56.6% of their day in general education, which are all considered acceptable. The CFI and NNFI were acceptable at .98 for all groups, but the SRMS ranged from .99 to .061 and was unacceptable for groups that had fewer than 150 students and for the group of students who spent more than 96.7% of their school day in general education. Fit statistics for this step can be found in the Test of Eight-Factor Structure Only column of Table 60.

Table 60

Model Fit and Change in Fit Statistics for TAGG-P by Percent of Time in General Education

Percent of Time in General Education		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Strength
<24.9%	χ2	3463.92	3501.26
n = 139	RMSEA	0.059	0.059
	SRMR	0.097	0.100
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.44	0.43
	ΔMc		0.00
	Gamma^	0.911	0.911
	ΔGamma^		0.000
25.0 – 39.2%	χ2	3596.94	3633.01
n = 114	RMSEA	0.061	0.060
	SRMR	0.098	0.120
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.42	0.42
	ΔMc		0.00
	Gamma^	0.907	0.906
	ΔGamma^		-0.001
39.3 – 53.6%	χ2	3447.01	3479.00
n = 140	RMSEA	0.057	0.056
	SRMR	0.098	0.110
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.44	0.44
	ΔMc		0.00
	Gamma^	0.912	0.911
	ΔGamma^		-0.001

Table 60

Model Fit and Change in Fit Statistics for TAGG-P by Percent of Time in General Education, continued

Percent of Time in General Education		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Strength
53.4 - 67.9%	$\chi 2$	3533.42	3562.43
n = 165	RMSEA	0.058	0.057
	SRMR	0.098	0.120
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.909	0.940
	ΔGamma^		0.031
68.0 - 82.3%	$\chi 2$	3542.77	3568.86
n = 302	RMSEA	0.059	0.058
	SRMR	0.078	0.083
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.42	0.42
	ΔMc		0.00
	Gamma^	0.908	0.908
	ΔGamma^		0.000
82.4 – 96.6%	$\chi 2$	3459.52	3494.01
n = 268	RMSEA	0.058	0.057
	SRMR	0.064	0.063
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.44	0.44
	ΔMc		0.00
	Gamma^	0.911	0.911
	ΔGamma^		0.000

Table 60

Model Fit and Change in Fit Statistics for TAGG-P by Percent of Time in General Education, continued

Percent of Time in General Education		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
96.7 – 100%	$\chi 2$	3463.92	3501.26
n = 320	RMSEA	0.059	0.059
	SRMR	0.097	0.100
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.44	0.44
	ΔMc		0.00
	Gamma^	0.912	0.912
	ΔGamma^		0.000

Strength of item-construct relations. No significant decline was found in the fit of the eight-factor structure when factor loadings from the full *TAGG* sample were applied to each of the subgroups, meaning educators believe the behaviors have the same values to their respective constructs regardless of how much a student participates in general education. For example, educators believed that a student telling someone about the supports he needs is equally valuable to having high Disability Awareness no matter how much time the student spends in general education.

Significant change was assessed using three change statistics—the ΔCFI, the ΔMc, and the ΔGamma^. Significance for these statistics is indicated when they are less than -.01, -.02, and -.001, respectively (Cheung & Rensvold, 2002). For this study, a significant decrease detected by only one change in fit statistic signified that the *TAGG* developers should use caution when suggesting goals. A significant decrease in two or more of the change in fit statistics signified the *TAGG* developers should examine the structure of the model for the group before they use the *TAGG* behaviors as suggested goals.

The ΔCFI and ΔMc were 0.00 for groups at all levels of participation in general education and ΔGamma^ ranged from -0.001 for groups who spent about a quarter to half of their day in general education to .031 for students who spent about 55-70% of their school day in general education, indicating the professional educators in this study did not view the *TAGG* behaviors as having greater or reduced value to the constructs depending on how much time students spend in general education. Fit statistics and change statistics can be found for each of the groups in Table 60 in the Test of Eight-Factor Structure and Item-Factor Strength column.

Transition education groups. I also wanted to understand whether exposure to transition education concepts such as learning Disability Awareness or how to actively participate or lead an IEP meeting contributed to the behaviors educators associated with the eight *TAGG-P* constructs, particularly because *TAGG* development targeted transition educators who likely already had knowledge of many of the skills the *TAGG* assesses, such as IEP participation.

Item-construct relations. This study found acceptable fit of the eight-factor structure for professional educators in the study whose students had little to no transition education (χ^2 = 2027.79, *df* = 1056, RMSEA = .053, CFI = .98, NNFI = .98, SRMR = .083) and for the professional educators in this study whose students had slightly more transition education (χ^2 = 2368.33, *df* = 1056, RMSEA = .06, CFI = .96, NNFI = .98, SRMR = .059), meaning educators believed the same behaviors are associated with the *TAGG* constructs even when a student has not had instruction in those areas. Fit statistics for this step in the analysis can be found in the Test of Eight-Factor Structure Only column of Table 61.

Table 61

Model Fit and Change in Fit Statistics for TAGG-P Level of Transition Education Instruction

Transition Education Instruction		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
Disability Awareness *or* IEP Involvement Instruction *or* none n = 333	$\chi2$	2027.79	2052.74
	RMSEA	0.053	0.052
	SRMR	0.083	0.064
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.52	0.52
	ΔMc		0.00
	Gamma^	0.928	0.928
	ΔGamma^		0.000
Disability Awareness *and* IEP Involvement Instruction n = 405	$\chi2$	2368.33	2391.44
	RMSEA	0.06	0.06
	SRMR	0.059	0.062
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.41	0.41
	ΔMc		0.00
	Gamma^	0.905	0.905
	ΔGamma^		0.000

Strength of item-construct relations. No significant decline in the fit of the *TAGG* structure was found when the factor loadings from the non-focus sample were applied to each of the transition education groups. This means professional educators in this study believed the behaviors were equally valuable to the measured constructs for students who had little to no instruction in transition education concepts as for students who had slightly more instruction in transition education concepts. The change in fit statistics are included in Table 61 in the Test of Eight-Factor Structure and Item-Factor Relations column.

Grade level groups. Because the *TAGG* is designed for all high school students with disabilities in any grade who plan on entering postsecondary education and competitive employment and because the developers believe the *TAGG* behaviors are learned and will improve with practice, I examined the suitability of the eight-construct *TAGG-P* structure for students in the 9^{th}-12^{th} grades, and 18-21 programs. While the developers understand and believe that the scores will be different for students at each grade level, it is hoped that professional educators believe the same behaviors are equally related to their constructs across grade levels.

Item-construct relations. This study found consistently good fit of the eight-factor *TAGG-P* model across grade levels. The RMSEAs for the grades ranged from .064 for the 18-21 year old program group to .058 for the 10^{th} grade group. The CFI and NNFI were .98 for all groups, but the SRMS ranged from .13 to .053 and appeared unacceptable for groups that had fewer participants. Fit statistics for this step can be found in the Test of Eight-Factor Structure and Item-Factor Relations column of Table 62.

Table 62

Model Fit and Change in Fit Statistics for TAGG-P by Grade Level

Grade		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
9th	$\chi2$	3669.67	3734.01
n = 124	RMSEA	0.06	0.06
	SRMR	0.13	0.16
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.406	0.402
	ΔMc		0.00
	Gamma^	0.904	0.903
	ΔGamma^		-0.001
10^{th}	$\chi2$	3471.95	3522.13
n = 162	RMSEA	0.058	0.058
	SRMR	0.12	0.13
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.910	0.910
	ΔGamma^		0.000
11^{th}	$\chi2$	3496.38	3527.1
n = 480	RMSEA	0.059	0.058
	SRMR	0.055	0.064
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.910	0.910
	ΔGamma^		0.000

Table 62

Model Fit and Change in Fit Statistics for TAGG-P by Grade Level, continued

Grade		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
12th	$\chi 2$	3640.37	3686.28
n = 628	RMSEA	0.059	0.059
	SRMR	0.054	0.055
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.41	0.41
	ΔMc		0.00
	Gamma^	0.905	0.905
	ΔGamma^		0.000
18-21 Program	$\chi 2$	3681.11	
n = 74	RMSEA	0.064	0.063
	SRMR	0.13	0.14
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		
	Mc	0.40	0.41
	ΔMc		0.00
	Gamma^	0.904	0.904
	ΔGamma^		0.000

Strength of item-construct relations. The ΔCFI, ΔMc, and ΔGamma^ were 0.0 for grades 10-12. The ΔCFI and ΔMc were both 0.0 for 9th grade students but the ΔGamma^ indicated a near significant decrease in model fit for 9th grade students. This means professional educators in this study may believe the same behaviors were associated with each of the *TAGG-P* constructs for 9th graders as for other high school students, but that the degree to which educators associate the behaviors with the constructs may be unique for 9th graders. This study does not examine the individual items and their relations to the construct, but the results for this analysis mean the professional educators in this study may, for example, believe the easier items are more suitable as goals for students in 9th and 10th grades than for students in 11th and 12th grades. Fit statistics and change statistics can be found for each of the groups in Table 62 in the Test of Eight-Factor Structure and Item-Factor Relations column.

Disability category. Educators consider their students' disability category when they select classroom materials and choose instructional methods for teaching. The *TAGG* developers wanted to respond to educators' desire for information specific to the categories of students they teach. The *TAGG-P* data included enough data to examine five groups of students with disabilities—students whose primary disabilities were autism, emotional disabilities, intellectual disabilities, other health impairments, or specific learning disabilities.

Item-construct relations. This study found consistently acceptable fit of the eight-factor *TAGG-P* model for each of the disability categories examined. The RMSEAs for the categories ranged from .069 for the ED group to .055 for the ID group. The CFI and NNFI were .98 for each of the groups except the ED group for which the CFI and NNFI were .97, but the SRMS ranged from .22 for the ED group to .06 for the SLD group. Overall, the fit was acceptable for all of the disability groups, but was slightly less acceptable for students having emotional disabilities. Fit statistics for this step can be found in the Test of Eight-Factor Structure Only column of Table 63.

Table 63

Model Fit and Change in Fit Statistics for TAGG-P by Disability Category

Disability Category		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
Autism	$\chi2$	3656.41	3696.29
n = 61	RMSEA	0.063	0.063
	SRMR	0.18	0.22
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.41	0.41
	ΔMc		0.00
	Gamma^	0.905	0.904
	ΔGamma^		0.000
ED	$\chi2$	4602.64	4665.69
n = 79	RMSEA	0.069	0.069
	SRMR	0.22	0.21
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.29	0.29
	ΔMc		0.00
	Gamma^	0.874	0.873
	ΔGamma^		-0.001
ID	$\chi2$	3115.01	3139.09
n = 174	RMSEA	0.055	0.054
	SRMR	0.11	0.12
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.49	0.49
	ΔMc		0.00
	Gamma^	0.923	0.922
	ΔGamma^		-0.001

Table 63

Model Fit and Change in Fit Statistics for TAGG-P by Disability Category, continued

Disability Category		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
OHI	χ2	3558.75	3415.59
n = 200	RMSEA	0.057	0.056
	SRMR	0.07	0.077
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.42	0.45
	ΔMc		0.03
	Gamma^	0.908	0.914
	ΔGamma^		0.006
SLD	χ2	3579.28	3415.59
n = 828	RMSEA	0.06	0.056
	SRMR	0.058	0.077
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.42	0.45
	ΔMc		0.03
	Gamma^	0.907	0.914
	ΔGamma^		0.007

Strength of item-construct relations. When relations of the items to the constructs was set equal for each group as for the non-focus group of participants, the fit of the eight-factor structure remained acceptable for each disability category group. The ΔCFI was 0.0 for each of the groups. The ΔMc ranged from 0.0 for the autism, ID, and ED groups to .03 for the OHI and SLD groups. The ΔGamma^ ranged from near significant for the ED and ID groups (ΔGamma^ = -.001) to .007 for the SLD group.

Fit statistics and change statistics can be found for each of the groups in Table 63 in the Test of Eight-Factor Structure and Item-Factor Relations column.

***TAGG* Family Version**

As with the *TAGG-P*, I wanted to understand whether the item-construct relations on the *TAGG-F* were suitable across groups of students who spent varying degrees of time in general education. The *TAGG-F* has shown acceptable fit in development (χ^2 = 862.74, *df* = 499, RMSEA = .0570, CFI = .91, TLI = .90, and RMSR = .058) and in replication (χ^2 = 1995.76, *df* = 1087, RMSEA = .0579, CFI = .89, TLI = .89, RMSR = .0679), though some differences have been observed in the strengths of the item-construct relations (Hennessey et al., 2013), which the developers attribute to families not having opportunities to observe students on some of the behaviors. The results presented below reflect the inclusion of all versions of the *TAGG-F* collected during the first three phases of development in the analysis.

Multigroup analysis. I used the same two-step process for the *TAGG-F* that I used for the *TAGG-P* to understand whether the eight-factor *TAGG* structure was appropriate for several groups of *TAGG* users. The first step of the analysis examined the test of the eight-factor configuration only while the second step looked at the configuration and the strengths of the relations of the items to their constructs.

Percent of time in general education groups. Groups for this analysis were the same as the groups for the analysis of the *TAGG-P*. The percent of time for each group and the number of participants in each group can be found in the Percent of Time in General Education Column of Table 64.

Item-construct relations. This study found consistently good fit of the eight-factor *TAGG-F* model across participation in general education, meaning family members associated the same behaviors with the same constructs for students regardless of their participation in general education. For example, no matter how much time their students spend in general education, families associated students using plans they made to attain their goals with the Goal Attainment construct.

The RMSEAs for the groups ranged from .059 for students who spent less than a quarter of their day in general education to .05 for students who spent more than 96% of their day in general education which are all considered acceptable. The NNFI was acceptable at .97 for groups that spent less than about 80% of their day in general education and increased to .98 for groups that spent more than about 80% of their day in general education. Similarly, the CFI was .97 for groups that spent up to about half their day in general education and improved to .98 for groups that spent more than half their day in general education. The SRMS ranged from .081 for students who spent about 70-80% of their school day in general education, but decreased to 0.19 for students who spent a quarter to 40% of their day in general education. Fit statistics for this step can be found in the Test of Eight-Factor Structure Only column of Table 64.

Table 64

Model Fit and Change in Fit Statistics for TAGG-F by Percent of Time in General Education

Percent of Time in General Education		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
<24.9%	$\chi 2$	2386.95	2422.09
n = 63	RMSEA	0.059	0.058
	SRMR	0.13	0.15
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.910	0.909
	ΔGamma^		-0.001
25.0 – 39.2%	$\chi 2$	2318.74	2345.65
n = 58	RMSEA	0.054	0.054
	SRMR	0.190	0.210
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.45	0.45
	ΔMc		0.00
	Gamma^	0.914	0.914
	ΔGamma^		0.000
39.3 – 53.6%	$\chi 2$	2332.96	2364.30
n = 76	RMSEA	0.055	0.055
	SRMR	0.160	0.170
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.44	0.44
	ΔMc		0.00
	Gamma^	0.912	0.912
	ΔGamma^		0.000

Table 64

Model Fit and Change in Fit Statistics for TAGG-F by Percent of Time in General Education, continued

Percent of Time in General Education		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
53.4 - 67.9%	$\chi 2$	2268.31	2281.47
n = 76	RMSEA	0.053	0.053
	SRMR	0.110	0.120
	NNFI	0.97	0.97
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.46	0.47
	ΔMc		0.00
	Gamma^	0.917	0.978
	ΔGamma^		0.061
68.0 - 82.3%	$\chi 2$	2229.51	2263.22
n = 168	RMSEA	0.054	0.053
	SRMR	0.081	0.083
	NNFI	0.97	0.97
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.47	0.47
	ΔMc		0.00
	Gamma^	0.919	0.919
	ΔGamma^		0.000
82.4 – 96.6%	$\chi 2$	2173.34	2182.00
n = 157	RMSEA	0.052	0.051
	SRMR	0.110	0.110
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.49	0.49
	ΔMc		0.01
	Gamma^	0.922	0.923
	ΔGamma^		0.001

Table 64

Model Fit and Change in Fit Statistics for TAGG-F by Percent of Time in General Education, continued

Percent of Time in General Education		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
96.7 – 100%	$\chi 2$	2103.26	2151.54
n = 159	RMSEA	0.050	0.051
	SRMR	0.090	0.100
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.51	0.50
	ΔMc		-0.01
	Gamma^	0.926	0.925
	ΔGamma^		-0.001

Strength of item-construct relations. No significant decline was found in the fit of the eight-factor structure when factor loadings from the full *TAGG* sample were applied to each of the subgroups, meaning families in this study believed the behaviors have the same values to their respective constructs regardless of how much a student participates in general education. For example, families in this study believed that a student knowing which assignments he will have trouble with when they are given is equally valuable to having high Knowledge of Strengths and Limitations no matter how much time the student spends in general education.

The ΔCFI was 0.00 for all of the groups. Similarly, the ΔMc was 0.00 for all groups, except for the group of student who spent about 80-95% of their school day in general education (ΔMc = 0.01). The ΔGamma^ was 0.00 for the groups of students who spent less than about 80% of their school day in general education, but was slightly lower for groups of students who spent more than about 80% of their day in general education. (ΔGamma^ = -0.001). Fit statistics and change statistics can be found for each of the groups in Table 64 in the Test of Eight-Factor Structure and Item-Factor Relations column.

Transition education groups. For consistency in the analysis across the *TAGG* versions, I divided the *TAGG-F* versions by those completed by family members of students who had exposure to concepts in transition education and students who had not. This allowed for an understanding of whether exposure to transition education contributed to the behaviors families associated with the eight *TAGG-F* constructs.

Item-construct relation. This study found acceptable fit of the eight-factor structure for families whose students had little to no transition education (χ^2 = 2075.68, *df* = 1056, RMSEA = .054, CFI = .98, NNFI = .98, SRMR = .065) and for families whose students had slightly more transition education (χ^2 = 2200.54, *df* = 1056, RMSEA = .053, CFI = .98, NNFI = .97, SRMR = .092), meaning families believed the same behaviors are associated with the *TAGG* constructs even when a student has not had instruction in the transition education concepts included in this study. Fit statistics for this step in the analysis can be found in the Test of Eight-Factor Structure Only column of Table 65.

Table 65

Model Fit and Change in Fit Statistics for TAGG-F Level of Transition Education Instruction

Transition Education Instruction		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
Disability Awareness *or* IEP Involvement Instruction *or* none n = 252	χ2	2175.68	2205.73
	RMSEA	0.054	0.053
	SRMR	0.065	0.069
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.49	0.48
	ΔMc		0.00
	Gamma^	0.922	0.921
	ΔGamma^		0.000
Disability Awareness *and* IEP Involvement Instruction n = 141	χ2	2175.68	2205.73
	RMSEA	0.054	0.053
	SRMR	0.065	0.069
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.49	0.48
	ΔMc		0.00
	Gamma^	0.922	0.921
	ΔGamma^		0.000

Strength of item-construct relations. No significant decline in the fit of the *TAGG* structure was found when the factor loadings from the full sample were applied to each of the transition education groups. This means families believed the behaviors were equally valuable to the measured constructs for students who had little to no instruction in transition education concepts. The change in fit statistics are included in Table 65 in the Test of Eight-Factor Structure and Item-Factor Relations column.

Grade level groups. To understand whether the families in this study may have different ideas about whether the *TAGG* behaviors are suitable for students at different grade levels, I examined the suitability of the eight-construct *TAGG-F* for students in the 9th-12th grades. Specifically, this analysis was designed to ensure that grade level does not need to be considered when goals are suggested on the *TAGG-F* scoring profile.

Item-construct relation. This study found consistently good fit of the eight-factor *TAGG-F* model across grade levels. The RMSEAs for the groups ranged from .057 for the 9th grade group to.053 for the 10th grade group. The NNFI was .97 for all groups, and the CFI was .97 for the 9th, 10th, and 12th grade groups, and 0.98 for the 11th grade group. The SRMS ranged from.11 for the 9th grade group to .067 for the 11th grade group. This suggests that families endorse students engaging in the behaviors included on the *TAGG-F* assessment whether the student is in the 9th grade or 12th grade. Fit statistics for this step can be found in the Test of Eight-Factor Structure Only column of Table 66.

Table 66

Model Fit and Change in Fit Statistics for TAGG-F by Grade Level

Grade		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
9th	$\chi 2$	2275.74	2333.01
n = 105	RMSEA	0.057	0.058
	SRMR	0.11	0.14
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.461	0.454
	ΔMc		-0.01
	Gamma^	0.917	0.915
	ΔGamma^		*-0.002*
10^{th}	$\chi 2$	2255.58	2280.51
n = 122	RMSEA	0.53	0.052
	SRMR	0.097	0.11
	NNFI	0.97	0.97
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.47	0.47
	ΔMc		0.00
	Gamma^	0.918	0.919
	ΔGamma^		0.001
11^{th}	$\chi 2$	2192.46	2211.92
n = 289	RMSEA	0.054	0.053
	SRMR	0.067	0.07
	NNFI	0.97	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.48	0.48
	ΔMc		0.00
	Gamma^	0.910	0.910
	ΔGamma^		0.000

Table 66

Model Fit and Change in Fit Statistics for TAGG-F by Grade Level, continued

Grade		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
12th	$\chi 2$	2273.95	2307.25
n = 241	RMSEA	0.054	0.054
	SRMR	0.083	0.09
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.45	0.45
	ΔMc		0.00
	Gamma^	0.915	0.915
	ΔGamma^		0.000

Strength of item-construct relations. The ΔCFI, ΔMc, and ΔGamma^ were non-significant for grades 10-12, but the fit of the eight-factor construct did decrease for students in the 9th grade, suggesting families in this study believe the same behaviors are appropriate for 9th graders as for students in the other grades, but that their perceptions of degree to which the behaviors are of value to the constructs is different for 9th grade students than for students in 10th-12th grades. This is most likely to be true for behaviors such as having a paid job or accessing social services, which families may not view as a high priority for students with four years of high school ahead. Fit statistics and change statistics can be found for each of the groups in Table 66 in the Test of Eight-Factor Structure and Item-Factor Relations column.

Disability category. Families may use disability category labels to understand the needs of their students, particularly when the students are attempting to access services through vocational rehabilitation or at a college disability services office.

Including disability categories allowed me to understand whether families in this study believe students with different disabilities need to be proficient in the same set of skills to score high on the *TAGG* constructs. Due to the number of family participants in the *TAGG* development, this study could only consider *TAGG-F* forms returned for student with emotional disabilities, other health impairments, and specific learning disabilities.

Item-construct relations. This study found consistently acceptable fit of the eight-factor *TAGG-F* model for the three disability categories examined. The RMSEAs for the categories ranged from .057 for the SLD group to .051 for the OHI group. The CFI and NNFI were .97 for the SLD and ED groups, and .98 for the OHI group. The SRMS ranged from .10 for the ED group to .051 for the OHI group. Overall, the fit was acceptable for all of the disability groups, indicating families in this study believed the same behaviors were associated with the same constructs for each disability category. Fit statistics for this step can be found in the Test of Eight-Factor Structure Only column of Table 67.

Table 67

Model Fit and Change in Fit Statistics for TAGG-F by Disability Category

Disability Category		Test of Eight-Factor Structure Only	Test of Eight-Factor Structure and Item-Factor Relations
ID	$\chi 2$	2350.48	2375.29
n = 97	RMSEA	0.056	0.055
	SRMR	0.10	0.11
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.44	0.44
	ΔMc		0.00
	Gamma^	0.912	0.910
	ΔGamma^		*-0.002*
OHI	$\chi 2$	2107.72	2122.32
n = 109	RMSEA	0.051	0.051
	SRMR	0.095	0.1
	NNFI	0.98	0.98
	CFI	0.98	0.98
	ΔCFI		0.00
	Mc	0.51	0.52
	ΔMc		0.01
	Gamma^	0.927	0.928
	ΔGamma^		0.001
SLD	$\chi 2$	2223.34	2280.56
n = 442	RMSEA	0.057	0.057
	SRMR	0.065	0.065
	NNFI	0.97	0.97
	CFI	0.97	0.97
	ΔCFI		0.00
	Mc	0.461	0.454
	ΔMc		-0.01
	Gamma^	0.917	0.915
	ΔGamma^		*-0.002*

Strength of item-construct relations. The ΔGamma^ indicated a significant decrease in model fit for the ED (ΔGamma^ = -.002) and SLD (ΔGamma^ = -.002) groups when the relations of the items to the constructs was set equal for each group as for the non-focus group of participants. This suggests family members in this study believed that the same behaviors were important for all three groups of students, but that they believed some behaviors may be of different priority for the SLD and ED groups. The ΔMc was 0.0 for the ED and OHI groups, but was -0.01 for the SLD group, giving some credence to the significance in the ΔGamma^ for the SLD group. The ΔCFI, however, was 0.0 for all three groups. Fit statistics and change statistics can be found for each of the groups in Table 67 in the Test of Eight-Factor Structure and Item-Factor Relations column.

TAGG **Student Version**

The *TAGG-S* was developed by applying the item-construct structure that emerged from initial analysis of the professional version. The eight-factor *TAGG-P* structure had less than acceptable fit for the student participant sample during development, and a review of the modification indices indicated the Knowledge of Strengths and Limitations construct and Community Supports construct be collapsed into one construct having eight items, rather than four per construct as in the professional and family versions (Hennessey et al., 2013).

Multigroup analysis. For this portion of the study, I used the same process I used to analyze the *TAGG-P* and *TAGG-F* assessments. First, I tested the fit of the seven-factor *TAGG-S* structure on each of the focus groups. Then, I constrained the item loadings of the focus group to be equal to the non-focus group to understand whether any groups of students had different beliefs about the value of the *TAGG-S* behaviors to their constructs.

Percent of time in general education groups. Students who participated in the analysis of the *TAGG-S* were grouped using the same methods as for the professional version, with the largest portion of the students spending more than half their school day in general education.

Item-construct relations. This study found consistently good fit of the seven-factor *TAGG-S* model across participation in general education, meaning students in this study associated the same behaviors with the same constructs regardless of their participation in general education. For example, no matter how much time students spent in general education, they associated telling their IEP team their plans for after high school with the Involvement in the IEP construct.

The RMSEAs for the groups ranged from .064 for students who spent less than a quarter of their day in general education to .056 for students who spent more than 96% of their day in general education. The NNFI ranged from .94-.95 for the groups and the CFI was .94 for all groups but those who spent more than 97% of their school day in general education. The SRMS ranged from .14 for students who spent a quarter to half of their day in general education to .074 for students who spent more than 96% of their day in general education. Overall, the fit of the seven-factor structure was acceptable for all of the groups, but the fit was remarkably better for groups of students who spent essentially all of their school day in general education. Fit statistics for this step can be found in the Test of Seven-Factor Structure Only column of Table 68.

Table 68

Model Fit and Change in Fit Statistics for TAGG-S by Percent of Time in General Education

Percent of Time in General Education		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
<24.9%	$\chi2$	3376.36	3415.56
n = 123	RMSEA	0.064	0.063
	SRMR	0.11	0.12
	NNFI	0.93	0.93
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.909	0.910
	ΔGamma^		0.001
25.0 – 39.2%	$\chi2$	3367.1	3411.6
n = 103	RMSEA	0.063	0.062
	SRMR	0.11	0.13
	NNFI	0.94	0.93
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.43	0.42
	ΔMc		-0.01
	Gamma^	0.909	0.908
	ΔGamma^		-0.001
39.3 – 53.6%	$\chi2$	3372.27	3405.66
n = 131	RMSEA	0.063	0.063
	SRMR	0.14	0.17
	NNFI	0.93	0.93
	CFI	0.94	0.93
	ΔCFI		-0.01
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.909	0.909
	ΔGamma^		0.000

Table 68

Model Fit and Change in Fit Statistics for TAGG-S by Percent of Time in General Education, continued

Percent of Time in General Education		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
53.4 - 67.9%	χ2	3309.92	3338.05
n = 157	RMSEA	0.062	0.061
	SRMR	0.088	0.097
	NNFI	0.93	0.94
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.44	0.44
	ΔMc		0.00
	Gamma^	0.911	0.911
	ΔGamma^		0.000
68.0 - 82.3%	χ2	3363.04	3387.71
n = 281	RMSEA	0.063	0.062
	SRMR	0.088	0.088
	NNFI	0.94	0.93
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.909	0.909
	ΔGamma^		0.000
82.4 – 96.6%	χ2	3346.88	3368.38
n = 255	RMSEA	0.061	0.06
	SRMR	0.092	0.094
	NNFI	0.93	0.93
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.910	0.910
	ΔGamma^		0.000

Table 68

Model Fit and Change in Fit Statistics for TAGG-S by Percent of Time in General Education, continued

Percent of Time in General Education		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
96.7 – 100%	$\chi2$	3314.1	2852.8
n = 302	RMSEA	0.056	0.055
	SRMR	0.074	0.073
	NNFI	0.94	0.95
	CFI	0.95	0.95
	ΔCFI		0.00
	Mc	0.44	0.52
	ΔMc		0.08
	Gamma^	0.911	0.929
	ΔGamma^		0.018

Strength of item-construct relations. No significant decline was found in the fit of the seven-factor structure when factor loadings from the non-focus *TAGG* groups were applied to each of the focus groups, meaning students who participated in this study believed the *TAGG-S* behaviors have the same values to their respective constructs regardless of how much a student participates in general education. For example, students believed that learning from mistakes is equally valuable to having high Persistence no matter how much time the student spends in general education.

The ΔCFI was 0.00 for most of the groups, but was near significant for the group of students who spent about a 40-50% their day in general education (ΔCFI = -.01). The ΔMc was 0.00 for all of the groups, except for students who spent less than a quarter of their day in general education (ΔMc = -.01). The ΔGamma^ ranged from -.001 for students who spent about 25-40% of their school day in general education to

.018 for students who spent more than 96% of their day in general education. Fit statistics and change statistics can be found for each of the groups in Table 68 in the Test of Seven-Factor Structure and Item-Factor Relations column.

Transition education groups. I wanted to understand whether exposure to certain transition education concepts contributed to the behaviors students associated with the seven *TAGG-S* constructs, since students who had transition education may have different beliefs about which items are associated with which constructs than students who have not had instruction in certain transition education concepts.

Item-construct relation. This study found acceptable fit of the seven-factor structure for students who had little to no instruction in transition education concepts (χ^2 = 3313.17, *df* = 1063, RMSEA = .076, CFI = .92, NNFI = .92, SRMR = .015) and for students who had instruction in some transition education concepts (χ^2 = 2921.77, *df* = 1063, RMSEA = .072, CFI = .92, NNFI = .92, SRMR = .081), meaning students who participated in this study believed the same behaviors were associated with the *TAGG* constructs even when they have not had instruction in those areas. Fit statistics for this step in the analysis can be found in the Test of Seven-Factor Structure Only column of Table 69.

Table 69

Model Fit and Change in Fit Statistics for TAGG-S Level of Transition Education Instruction

Transition Education Instruction		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
Disability Awareness *or* IEP Involvement Instruction *or* none n = 315	$\chi 2$	3313.17	2847.1
	RMSEA	0.076	0.075
	SRMR	0.15	0.17
	NNFI	0.92	0.91
	CFI	0.92	0.92
	ΔCFI		0.00
	Mc	0.22	0.22
	ΔMc		0.00
	Gamma^	0.848	0.849
	ΔGamma^		0.001
Disability Awareness *and* IEP Involvement Instruction n = 395	$\chi 2$	2921.77	2981.88
	RMSEA	0.072	0.073
	SRMR	0.081	0.085
	NNFI	0.90	0.90
	CFI	0.90	0.90
	ΔCFI		0.00
	Mc	0.283	0.277
	ΔMc		-0.006
	Gamma^	0.871	0.869
	ΔGamma^		*-0.002*

Strength of item-construct relations. No significant decline in the fit of the seven-factor *TAGG* structure was found when the factor loading from the non-focus groups were applied to each of the transition education groups. This means students believed the behaviors were equally valuable to the measured constructs whether they had little to no instruction in transition education concepts, or some instruction in transition education concepts. The change in fit statistics are included in Table 69 in the Test of Seven-Factor Structure and Item-Factor Relations column.

Grade level groups. Students are exposed to the skills assessed on the *TAGG-S* across their time in high school. To understand whether the length of time students have been exposed to the skills contributes to whether they believe the same items are related to the same constructs on the *TAGG-S*, the fit of the seven-factor structure was examined across grade levels.

Item-construct relations. This study found consistently acceptable fit of the seven-factor *TAGG-S* model across grade levels. The RMSEAs for the grades ranged from .065 for 9th grade students to .062 for students in the 10th and 11th grades. The NNFI was .93 for all groups, and the CFI was .93 for the 9th, 12th and 18-21 year old students and .94 for students in the 10th and 11th grades. The SMRS ranged from .18 for students in 18-21 year old programs to .75 for students in the 12th grade. Fit statistics for this step can be found in the Test of Seven-Factor Structure Only column of Table 70.

Table 70

Model Fit and Change in Fit Statistics for TAGG-S by Grade Level

Grade		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
9th	$\chi 2$	3489.00	3520.15
n = 117	RMSEA	0.065	0.064
	SRMR	0.15	0.16
	NNFI	0.93	0.93
	CFI	0.93	0.93
	ΔCFI		0.00
	Mc	0.41	0.41
	ΔMc		0.00
	Gamma^	0.905	0.906
	ΔGamma^		0.001

Table 70

Model Fit and Change in Fit Statistics for TAGG-S by Grade Level, continued

Grade		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
10th	$\chi 2$	3303.71	3321.21
n = 152	RMSEA	0.062	0.061
	SRMR	0.10	0.11
	NNFI	0.93	0.94
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.44	0.44
	ΔMc		0.00
	Gamma^	0.911	0.912
	ΔGamma^		0.001
11th	$\chi 2$	3335.95	3368.62
n = 429	RMSEA	0.062	0.061
	SRMR	0.08	0.081
	NNFI	0.93	0.94
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.43	0.43
	ΔMc		0.00
	Gamma^	0.910	0.910
	ΔGamma^		0.000
12th	$\chi 2$	3567.14	3616.07
n = 587	RMSEA	0.064	0.064
	SRMR	0.075	0.076
	NNFI	0.93	0.93
	CFI	0.93	0.93
	ΔCFI		0.00
	Mc	0.40	0.39
	ΔMc		0.00
	Gamma^	0.902	0.901
	ΔGamma^		-0.001

Table 70

Model Fit and Change in Fit Statistics for TAGG-S by Grade Level, continued

Grade		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
18-21 Program	$\chi 2$	3605.74	3629.77
n = 67	RMSEA	0.063	0.062
	SRMR	0.18	0.18
	NNFI	0.93	0.93
	CFI	0.93	0.93
	ΔCFI		0.00
	Mc	0.39	0.39
	ΔMc		0.00
	Gamma^	0.901	0.901
	ΔGamma^		0.000

Strength of item-construct relations. The ΔCFI and ΔMc were 0.00 for all grade levels. The ΔGamma^ ranged from -.001 for students in the 12th grade, which was near significant, to .001 for students in the 9th grade. This means students in this study generally believed the same behaviors were associated with each of the *TAGG-S* constructs and that the degree to which students associate the behaviors with the constructs did not change as a function of grade level. Fit statistics and change statistics can be found for each of the groups in Table 70 in the Test of Seven-Factor Structure and Item-Factor Relations column.

Disability category. Including the voice of the student in the IEP means using assessments that reflect student perspectives. Although this study takes the perspective that disability category likely does not contribute to *TAGG-S* scores, disability groups were analyzed to test this perspective.

Item-construct relation. Fit of the seven-factor *TAGG-S* structure was acceptable for each of the disability categories examined. The RMSEAs for the categories ranged from .073 for the ED group to .061 for the OHI group. The NNFI ranged from .89 for the ED group to .94 for the ID and OHI groups and the CFI ranged from .90 for the ED group to .94 for the ID, OHI, and SLD groups. Overall, the fit was acceptable for all of the disability groups, but as was the case with the *TAGG-P*, the fit was slightly less acceptable for students having emotional disabilities. Fit statistics for this step can be found in the Test of Seven-Factor Structure Only column of Table 71.

Table 71

Model Fit and Change in Fit Statistics for TAGG-S by Disability Category

Disability Category		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
Autism	$\chi 2$	3550.55	3588.78
n = 55	RMSEA	0.064	0.063
	SRMR	0.20	0.22
	NNFI	0.93	0.93
	CFI	0.93	0.93
	ΔCFI		0.00
	Mc	0.40	0.40
	ΔMc		0.00
	Gamma^	0.903	0.903
	ΔGamma^		0.000

Table 71

Model Fit and Change in Fit Statistics for TAGG-S by Disability Category, continued

Disability Category		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
ED	χ2	4749.34	4797.61
n = 71	RMSEA	0.073	0.073
	SRMR	0.13	0.13
	NNFI	0.89	0.89
	CFI	0.90	0.90
	ΔCFI		0.00
	Mc	0.26	0.25
	ΔMc		-0.01
	Gamma^	0.862	0.861
	ΔGamma^		-0.001
ID	χ2	3150.16	3170.16
n = 154	RMSEA	0.062	0.061
	SRMR	0.12	0.13
	NNFI	0.94	0.94
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.46	0.46
	ΔMc		0.00
	Gamma^	0.917	0.917
	ΔGamma^		0.000
OHI	χ2	3202.3	3244.13
n = 189	RMSEA	0.061	0.06
	SRMR	0.1	0.11
	NNFI	0.94	0.94
	CFI	0.94	0.94
	ΔCFI		0.00
	Mc	0.45	0.45
	ΔMc		0.00
	Gamma^	0.915	0.914
	ΔGamma^		-0.001

Table 71

Model Fit and Change in Fit Statistics for TAGG-S by Disability Category, continued

Disability Category		Test of Seven-Factor Structure Only	Test of Seven-Factor Structure and Item-Factor Relations
SLD	$\chi 2$	3393.44	3437.63
n = 778	RMSEA	0.062	0.062
	SRMR	0.073	0.072
	NNFI	0.93	0.93
	CFI	0.94	0.93
	ΔCFI		-0.01
	Mc	0.42	0.42
	ΔMc		0.00
	Gamma^	0.908	0.908
	ΔGamma^		0.000

Strength of item-construct relations. When relations of the items to the constructs was set equal for each disability category group as for the other participants, the fit of the seven-factor structure remained acceptable for each disability category group. The ΔCFI was 0.0 for all of the groups but the SLD group (ΔCFI = -.01). The ΔMc was 0.00 for all of the groups, except the group of students having ED (ΔMc = -.01). The ΔGamma^ was 0.000 for the autism, SLD, and ID groups, and -.001 for the SLD, OHI, and ED groups. Fit statistics and change statistics can be found for each of the groups in Table 71 in the Test of Seven-Factor Structure and Item-Factor Relations column.

The results presented in this chapter are positive for all three versions of the *TAGG* and indicate that users of the *TAGG* associate the same items with the same constructs regardless of the amount of time the students spend in general education. Across the board, professional educators, students, and students' families associated the same actions with the same constructs. For example, *TAGG* users consistently associated not giving up with high levels Persistence.

In terms of *TAGG* development, the findings are positive for two reasons. First, the findings of this study provide evidence that the *TAGG* is a suitable assessment for students who participate in general education to any degree, across several disabilities, and across several grade levels. It appears, based on this study, that *TAGG* scores should not increase or decrease as a direct result of general education participation, grade level, or exposure to concepts in transition education. Instead, it can be expected that as students have more exposure to the concepts presented in the *TAGG*, their ability with the *TAGG* behaviors will improve, thereby causing a score increase. This is ideal when measuring a student's ability.

Second, these findings provide evidence that using *TAGG* construct scores to generate goals for the broad range of students the *TAGG* targets is indeed a valid use of test scores. Prior to this study, it was unknown whether professional educators, families, and students would believe they should expect the same behaviors for students across their participation in general education, disability category, and grade level. This study found that the *TAGG* construct definitions (and thus, the goals generated from the definitions) were appropriate across each of those grouping variables.

CHAPTER FIVE

Discussion

Educators are required by federal law to include transition assessment data in the annual transition plans they develop for their students' IEPs. It is expected that because data is used to develop the plans, the plans will be of better quality, thereby leading to positive postsecondary outcomes as defined by Indicator 14—specifically, enrollment in postsecondary education and competitive employment. This may be true if the data used is technically sound and comprehensive, which, unfortunately, is not the case for many transition assessments (Tables 3 and 4). Although most of the commonly used transition assessments provide evidence of their technical properties, they do little to address using their test scores. Furthermore, it appears that the impact of disability is rarely considered in the development of transition assessments (Tables 5 and 6), even though the *Standards For Educational and Psychological Testing* (AERA et al., 1999) advise doing so.

Disability and accommodation have likely been neglected for two reasons—one practical and one methodological. First, obtaining enough participants with disabilities can be a challenge, particularly if the purpose of collecting participants is to group them by disability category, which is a necessary step for selling a test because test users want to know how the test works for students in a given disability category. To satisfy this aspect of product promotion, test developers need to define which students with disabilities a test is for and risk losing test users if they do not endorse use of the test for certain groups of students with disabilities. Second, because obtaining a large number of participants in specific disability categories is a challenge, test developers experience restrictions on how they can approach understanding the impact of disability and accommodations on test scores. The methods used for addressing other student characteristics—such as gender or ethnicity—require a large sample, especially if the goal is to detect individual items whose scores are unintentionally impacted by an unintended characteristic (i.e. gender, ethnicity, or disability).

However, the practical and methodological issues associated with understanding the impact of disability and accommodations do not excuse test developers from addressing the issue. After all, having a disability extends beyond having a categorical label and it may be argued that having a disability shapes students' academic experiences far more than other demographic variables. In some instances, students must spend a greater than ideal portion of their day in a restricted setting to have access to a special educator, while in other cases, students must spend all of their day in a less restrictive setting with limited access to a special educator. Understanding issues like these can help test developers find creative ways to conceptualize disability that extend beyond disability category.

Part of re-conceptualizing disability means understanding where it fits within a test's validity and reliability. Disabilities themselves are personal characteristics that may unintentionally influence test scores, indicating they may be a source of test bias. It is easy to understand that if a student with a math calculation disability is given an advanced math test, his ability to provide a correct response is impacted by his math calculation disability, not his ability at following the steps to answer an advanced math problem. The solution to the test bias created from by the student's disability is typically to provide the student with a calculator, but doing so involves violating the standardized administration procedures designed to reduce the outside factors that can contribute to test scores. Clearly, disabilities and accommodations present formidable challenges to precisely measuring students' ability, and the solution of providing an accommodation can further confound attempts to assess students.

Transition Assessment Test Scores

It seems that transition assessments serve a very different purpose from academic assessments, primarily because the scores from transition assessments have different use than scores from academic assessments. Academic test scores are used to measure growth or understand how a student is doing compared to his peers. The purpose of providing accommodations is to eliminate the impact of disability, thereby producing the highest scores a student can attain. In contrast, transition assessments are not interested in how many skills a student can do, and since goals will be created using the scores they generate, transition assessments are more concerned with precisely identifying specific skills a student can or cannot do.

Because the purpose of transition assessments is to identify behaviors rather than obtaining the highest score possible, transition assessment developers should not treat disabilities as a factor that impedes an overall test score the way developers of academic assessments would. Rather, transition assessment developers need to find ways to address the intersection of disabilities and the specific skills they intend to assess. For the most part, transition assessment developers have done this by developing assessments of skills most educators would want for any of their students, targeting the assessments for use with students with disabilities, and providing norms for students with disabilities.

For example, the *Transition Planning Inventory* (*TPI*; Clark & Patton, 1989*)*, which in a review of the commonly used transition assessments was the most commonly used transition assessment marketed for use with students with disabilities, assesses students with disabilities' knowledge of how to find a job, how to manage a home, and level of self-determination, skills and characteristics that are not unique to students with disabilities. The norms provided in the *TPI* reflect students with disabilities, which can help educators who use the *TPI* to understand how their students are doing compared to other students with disabilities, but this approach does not address the impact of disability and accommodations on test scores.

Ultimately, the most meaningful "test scores" obtained from transition assessments have nothing to do with the overall score, the most valuable transition assessment results come from the behaviors identified as weaknesses that can be turned into goals. That said, understanding the impact of disability and accommodations on transition assessments involves examining the relation between disability and the behaviors assessed by the instrument, not the relation between disability and the observed test score (i.e. level of self-determination, transition preparedness, etc.).

Re-conceptualizing Disability for Transition Assessment

Simply designating a test for use with students with disabilities and providing norms for students with disabilities and across disability categories does not meet the needs of transition assessment development because it does not improve the understanding of how disability relates to the behaviors included on a test. The traditional method for validating the behaviors included on an assessment is factor analysis, and because, once again, transition assessments are most concerned with behaviors, it is natural for factor analysis to continue to be a part of the validity evidence included in their technical materials.

However, these methods require sample sizes that are considered quite large by the special education research community. For instance, if a transition assessment developer wanted to provide evidence of the validity of the behaviors on his assessment for high school students with the Traumatic Brain Injury (TBI) categorical label, he could undertake a crude factor analysis that seeks only to confirm the configuration of the behaviors and constructs (Marsh & Balla, 1994). To do this, he would need to collect a minimum of 150 high school students with TBI who were able to take the assessment and for whom he could obtain parent consent. Identifying and obtaining consent for that many students alone is an ambitious goal, but the developer is also challenged with explaining how other factors, such as grade and gender, impact the performance of this small sample on his test. A test developer is likely to experience these challenges for any of the low incidence disability populations.

Clearly, disability category presents practical problems that limit test developers from using them to understand the relation between disability and the behaviors included on assessments, and transition assessment developers need to find new ways to understand the essence of disability. The primary focus of this study was to look at how disability impacts students' exposure to the behaviors included on a transition assessment by considering the percent of the school day students spend in general education—a factor driven by having a disability, but not the disability itself.

TAGG Validity Evidence

The transition assessments reviewed for the development of this project generally encouraged educators to develop goals by finding the assessment items on which students scored low, and then turning those items into annual transition goals. The *TAGG* is unique among transition assessments because it bridges the behaviors include on the assessment with transition planning by generating annual transition goals as part of the scoring profile. This is both convenient and ensures that annual transition goals meet the criteria for good goal writing (i.e. the goal is measurable and specific, with a criterion level that can be used to assess whether the goal was met). The *TAGG* developers have been seeking a way to understand whether the goals the *TAGG* generate are valid, but prior to this study, they had not developed a formal plan undertaking this task, most likely because the goal generation component of the *TAGG* is novel, and no examples exist for providing validity evidence of goals generated directly from an assessment.

To develop a method for assessing the validity of the goals generated by the *TAGG*, it is important to understand the four different roles of the behaviors included on the *TAGG*. First, the most obvious role of the behavior is that of *TAGG* assessment items. Each *TAGG* item assesses a specific behavior identified to be associated with postsecondary education and competitive employment in the transition research (McConnell et al., 2013). Second, each *TAGG* behavior is also a portion of one of the eight *TAGG* construct definitions, which are essentially lists of the behaviors associated with high scores on the construct (Hennessey et al., 2013; Martin et al., 2011). Third, the behaviors included on the *TAGG* serve as the goals suggested on the *TAGG-P*, *TAGG-F*, and *TAGG-S* score reports. Finally, when the *TAGG* behaviors are adopted as goals into transition plans, they become expectations.

The *TAGG* behaviors as expectations is particularly valuable given that the *TAGG* framework is being promoted as a new set of skills students should learn if they, their families and their educators want the students to attend postsecondary education and enter competitive employment. By pushing this framework, the *TAGG* developers send a message that students who exhibit the *TAGG* behaviors are more likely to attend postsecondary education and enter competitive employment, and thus, it is fair for educators and families to expect these behaviors from those students. Limiting the meaning of any analysis of the behaviors of the *TAGG* to simply a test of the construct definitions neglects the implications associated with the four roles of the *TAGG* behaviors. The *TAGG* is designed specifically to generate goals based on its construct definitions so any test of its construct definitions is a test of all the roles of the *TAGG* behaviors.

Disability and the *TAGG*

For this study, testing the validity of the *TAGG* goals means understanding whether it is acceptable for the *TAGG-P*, *TAGG-F*, and *TAGG-S* to generate the same goals across a disability related variable—participation in general education. To do this, I approached the construct definitions with an understanding that testing the validity of the construct definitions is the same as testing the validity of the goals. By this rational, if the construct definitions are suitable across the various groups of participants included in this study, then it can be assumed that the goals generated from those construct definitions are also valid.

From a measurement perspective, I wanted to understand whether participation in general education was an unintended factor associated with the construct definitions (and expectations) for students who spent more or less time in general education. To do so, I used a multigroup confirmatory factor analysis to understand whether the behaviors the *TAGG* developers are promoting were the same for students who spent most, some, or very little of their day in general education. To gain a complete picture of the role exposure to the *TAGG* skills may play in *TAGG* scores, the study was expanded to consider other ways students may experience the *TAGG* framework—such as progress through high school and having had instruction in transition education concepts such as Disability Awareness and Involvement in the IEP.

The volume of information presented in the results of this study extends beyond that which is typically found in a multigroup analysis, largely due to the way in which the groups were divided for the percent of time in general education study component. Substantial discussion was held to design the method for developing these groups, particularly because percent of time in general education was originally designed to be an indicator of level of need rather than a measure of exposure to the *TAGG* skills. For example, when students attend a full-inclusion style school, the percent of time they spend in general education would not be an indicator of their level of need. Ultimately, it was decided that while the percent of the school day was not an indicator of level of disability *for this study*, percent of time in general education could, and should, be retained as a grouping variable because of the strong case made around its potential impact on *TAGG* scores.

Once this was established, decisions needed to be made about how the sample would be divided. This was particularly challenging for two reasons. First, no study using this variable could be located to present as a guideline for dividing the groups. Second, this study wanted to produce a series of deeply meaningful results for a continuous variable, which necessitated going beyond simply comparing the suitability of the *TAGG* framework for two arbitrarily divided groups.

To resolve these issues, the study needed to balance creating the maximum number of groups with the minimum number of participants per group. This could have been done by dividing the sample into groups of about 100 students and letting their levels of participation in general education define the groups, but instead, I decided to let the normal distribution define the groups and used half-standard deviation units to set the minimum and maximum levels of general education participation for the group. Following this method allowed for an understanding of each group's participation in general education relative to other students within the study.

TAGG Development Findings

This study succeeded in answering the two research questions for all three versions of the *TAGG*. In response to research question one (Are the items on the *TAGG* associated with the same constructs regardless of group membership?), it was established that the *TAGG-P, TAGG-F,* and *TAGG-S* structures are indeed appropriate for students regardless of their participation in general education, grade level, exposure to transition education, and disability category. It appears educators, families, and students in this study believe that the behaviors assessed by the *TAGG* are appropriate indicators of the *TAGG* constructs for students with disabilities who plan on entering postsecondary education or competitive employment who spend all, part, or none of their school day in general education, in any high school grade, and whether the students had some or no transition education.

For the *TAGG* developers, this means the goals generated from the *TAGG*, which are based on the behaviors included in the *TAGG,* do not need to be adjusted based on a student's participation in general education, grade, exposure to some transition education, or (for the most part) disability category. Valid use of test scores means, in part, that test scores are only used for the purpose for which they are intended. For the *TAGG* developers, the *TAGG* scores are used to measure strengths and weaknesses, and create goals related to a specific set of behaviors. Had these behaviors not been appropriate to any of the groups, not only would the scores be inaccurate, but their use for generating goals would have been a huge violation of the validity of the *TAGG-P*, *TAGG-F*, and *TAGG -S* score reports. This would have left a student working on a behavior or skill for a full year of school that did not necessarily reflect his needs.

This study was also able to answer the second research question (Do the items on the *TAGG* make the same contribution to *TAGG* scores regardless of group membership?) for the *TAGG –P*, *TAGG-F*, and *TAGG-S*. For each of the groups, it appears that educators, families, and students in this study believed the *TAGG* behaviors are of equal value to the constructs they measure for most of the groups examined, the exceptions being students with Specific Learning Disabilities (SLD) and Intellectual Disabilities (ID) on the *TAGG-F* and students having disability awareness and instruction in leading and/or actively participating in their IEP meetings on the *TAGG-*

S. Among the groups that exhibited no significant change in the model fit, when educators completed *TAGG-P*s for students who spent little to no time in general education, for example, they associated changing goal plans that did not work with goal attainment to the same degree as for students who spent most of their time in general education.

On the other hand, families seemed to believe that some of the behaviors included on the *TAGG* were of more or less value to students having SLD or ID than other *TAGG* participants. This difference may be related to the postsecondary plans families have for their students—families of students with intellectual disabilities may not view their children as headed toward postsecondary education, and thus may prioritize behaviors associated with competitive employment. Similarly, families of students with SLD may view their children as moving toward postsecondary education, but without employment, with an understanding that their children cannot balance being enrolled in school and having a job. Among students, having transition education seemed to cause a decrease in the fit of the *TAGG* structure, suggesting that some of the behaviors on the *TAGG* have a different priority when students have exposure to instruction in Disability Awareness and instruction in leading and/or actively participating in their IEP meetings.

This portion of the findings was somewhat surprising as I expected that teachers would believe some of the *TAGG* behaviors were of lesser value than others depending on a student's grade or level of need. For example, it was unanticipated that educators would believe that leading an IEP meeting or accessing social services agencies would be equally valuable to such a broad range of students when measuring Involvement in the IEP or Community Supports. Referring back to the example of Raoul and Oscar, who in the introduction to this study received special education services in the form of speech articulation classes and full day special education support, respectively, it was surprising that educators seem to believe that Disability Awareness for Raoul and Oscar meant talking to people about their accommodations and telling others why they received special education services. After all, if Raoul does not believe he needs to talk about his disability and his disability is not a large part of his life, does he have poor disability awareness? Likewise, if Oscar has a disability that impacts his school placement and future employment opportunities, it was surprising that educators would endorse him thinking of his disabilities as only one facet of his life when it has a more global effect.

Implications for *TAGG* development. Taken together, these findings presented are, for the most part, positive for the *TAGG* developers. Strong evidence has been established that ensures the scores used from the *TAGG-P*, *TAGG-F*, and *TAGG-S* are being used appropriately and validly for suggesting annual transition goals for students' IEP transition plans. As it stands, goals will not be suggested that present a conflict for students, families, and educators during transition plan development, but ultimately, because families are responsible for ensuring that many of the nonacademic goals included on the *TAGG* are attained, educators should defer to them when selecting goals.

Peripheral Findings

The design of this study did not indicate using the factor loadings and R^2 values obtained during analysis to consider the specific relations between the *TAGG* items and their constructs, even though those relations are typically examined in a traditional CFA analysis.

For clarification, factor loadings are an indication of contribution a construct has to a test score. For example, a *TAGG-P* Goal Attainment item about moving onto the next goal when one is met has a factor loading of about .87 for students across all grade levels. Therefore, when a student responds to the item, 87% of the reason he gave his response can be attributed to his Goal Attainment ability. Factor loadings of |.35-.40| are generally viewed as acceptable (Tabachnick & Fidell, 2007) so that 35-40% of an item's score can be explained by the construct the item measures. R^2 values represent the percent contribution an item makes to a score. For example, an item having an R^2 of .20 makes up 20% of a student's score. Essentially, R^2 values reveal how much an item changes the total score. R^2 values of .15 or greater are generally viewed as acceptable (Tabachnick & Fidell, 2007), meaning each of the items on an assessment make up at least 15% of the test taker's score. The factor loadings and R^2 values for each of the groups (percent of time in general education, exposure to transition education concepts, grade level, and disability category) across each of the three *TAGG* versions are presented in tables 72-83 below.

Table 72

TAGG-P Factor Loadings and R^2 values by Percent of Time in General Education

	Factor Loading							R^2 Value						
Item	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100
sl1	0.72	0.72	0.72	0.71	0.72	0.72	0.72	0.53	0.52	0.52	0.51	0.51	0.52	0.51
sl2	0.83	0.80	0.68	0.70	0.77	0.78	0.75	0.62	0.60	0.52	0.53	0.58	0.59	0.56
sl3	0.77	0.76	0.73	0.78	0.77	0.74	0.80	0.62	0.62	0.59	0.64	0.62	0.60	0.68
sl4	0.92	0.85	0.77	0.68	0.73	0.71	0.76	0.68	0.64	0.59	0.53	0.56	0.54	0.58
da1	0.68	0.69	0.69	0.70	0.69	0.69	0.70	0.49	0.48	0.47	0.49	0.47	0.48	0.49
da2	0.92	0.87	0.81	0.75	0.79	0.78	0.81	0.73	0.70	0.67	0.63	0.66	0.66	0.67
da3	0.63	0.70	0.67	0.67	0.54	0.47	0.62	0.39	0.45	0.43	0.43	0.32	0.26	0.49
da4	0.71	0.66	0.62	0.54	0.45	0.61	0.55	0.43	0.40	0.37	0.60	0.23	0.35	0.28
p1	0.80	0.79	0.79	0.80	0.80	0.79	0.79	0.64	0.63	0.62	0.64	0.64	0.63	0.63
p2	0.90	0.88	0.93	0.87	0.91	0.92	0.91	0.82	0.81	0.83	0.81	0.82	0.83	0.82
p3	0.82	0.83	0.88	0.85	0.91	0.89	0.88	0.77	0.77	0.80	0.78	0.80	0.80	0.81
p4	0.90	0.84	0.96	0.86	0.91	0.92	0.91	0.82	0.80	0.85	0.81	0.83	0.83	0.84
p5	0.77	0.80	0.85	0.84	0.84	0.82	0.82	0.65	0.66	0.69	0.69	0.69	0.67	0.68
io1	0.78	0.79	0.78	0.79	0.79	0.79	0.79	0.61	0.63	0.61	0.62	0.62	0.63	0.62
io2	0.27	0.39	0.35	0.46	0.55	0.61	0.55	0.09	0.17	0.15	0.23	0.30	0.35	0.38
io3	0.80	0.84	0.88	0.60	0.83	0.70	0.79	0.62	0.65	0.67	0.48	0.64	0.55	0.62
ga1	0.78	0.79	0.79	0.79	0.78	0.78	0.78	0.61	0.62	0.63	0.62	0.61	0.61	0.60
ga2	0.93	0.80	0.94	0.88	0.88	0.92	0.90	0.82	0.78	0.83	0.80	0.81	0.82	0.83
ga3	0.97	0.89	0.89	0.88	0.93	0.93	0.93	0.89	0.86	0.87	0.86	0.87	0.88	0.88
ga4	0.92	0.80	0.84	0.82	0.92	0.94	0.89	0.82	0.77	0.79	0.78	0.82	0.83	0.82

Table 72

TAGG-P Factor Loadings and R^2 values by Percent of Time in General Education, continued

	Factor Loadings							R^2 Values						
Item	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100
ga5	0.94	0.77	0.82	0.86	0.87	0.86	0.86	0.78	0.70	0.72	0.75	0.75	0.75	0.77
ga6	0.34	0.44	0.56	0.29	0.28	0.29	0.34	0.16	0.18	0.28	0.08	0.08	0.09	0.11
e1	0.80	0.41	0.81	0.81	0.80	0.81	0.82	0.65	0.65	0.66	0.66	0.63	0.65	0.67
e2	0.88	0.89	0.93	0.88	0.91	0.85	0.89	0.79	0.78	0.80	0.79	0.80	0.79	0.78
e3	0.30	0.35	0.34	0.10	0.28	0.21	0.27	0.09	0.11	0.11	0.10	0.08	0.05	0.08
e4	0.41	0.41	0.43	0.34	0.30	0.43	0.41	0.16	0.17	0.18	0.12	0.10	0.18	0.20
iep1	0.69	0.69	0.70	0.69	0.69	0.69	0.70	0.48	0.47	0.49	0.48	0.48	0.47	0.49
iep2	0.90	0.88	0.85	0.90	0.96	0.93	0.91	0.83	0.81	0.81	0.83	0.82	0.82	0.82
iep3	0.90	0.75	0.87	0.98	0.96	0.92	0.91	0.82	0.76	0.82	0.83	0.84	0.82	0.82
iep4	0.64	0.56	0.53	0.62	0.66	0.65	0.61	0.40	0.33	0.31	0.39	0.41	0.41	0.35
cs1	0.72	0.7	0.71	0.72	0.70	0.72	0.70	0.51	0.51	0.51	0.51	0.52	0.51	0.52
cs2	0.62	0.58	0.59	0.67	0.65	0.54	0.60	0.38	0.35	0.35	0.43	0.40	0.32	0.41
cs3	0.75	0.57	0.75	0.72	0.75	0.70	0.74	0.54	0.41	0.54	0.53	0.55	0.52	0.53
cs4	0.53	0.55	0.41	0.28	0.37	0.34	0.36	0.25	0.26	0.16	0.08	0.14	0.12	0.13

Table 73

TAGG-P Factor Loadings and R^2 values by Transition Education

	Factor Loading		R^2 Value	
Item	Little to no Transition Education	Some Transition Education	Little to no Transition Education	Some Transition Education
sl1	0.73	0.71	0.53	0.51
sl2	0.77	0.74	0.58	0.55
sl3	0.77	0.78	0.62	0.63
sl4	0.80	0.73	0.61	0.56
da1	0.68	0.68	0.46	0.47
da2	0.86	0.74	0.70	0.62
da3	0.64	0.53	0.4.	0.32
da4	0.53	0.52	0.29	0.28
p1	0.80	0.79	0.64	0.62
p2	0.91	0.92	0.82	0.83
p3	0.90	0.88	0.80	0.79
p4	0.93	0.92	0.83	0.83
p5	0.82	0.82	0.67	0.68
io1	0.79	0.78	0.62	0.61
io2	0.50	0.49	0.26	0.25
io3	0.72	0.81	0.56	0.64
ga1	0.78	0.79	0.61	0.61
ga2	0.90	0.92	0.82	0.82
ga3	0.92	0.91	0.87	0.87
ga4	0.88	0.92	0.80	0.82

Table 73

TAGG-P Factor Loadings and R^2 values by Transition Education, continued

	Factor Loading		R^2 Value	
Item	Little to no Transition Education	Some Transition Education	Little to no Transition Education	Some Transition Education
ga5	0.90	0.86	0.77	0.74
ga6	0.27	0.32	0.08	0.10
e1	0.80	0.81	0.64	0.65
e2	0.90	0.89	0.8	0.78
e3	0.24	0.13	0.06	0.02
e4	0.41	0.37	0.17	0.14
iep1	0.70	0.69	0.49	0.48
iep2	0.91	0.91	0.83	0.83
iep3	0.91	0.94	0.83	0.83
iep4	0.65	0.63	0.41	0.38
cs1	0.72	0.72	0.52	0.51
cs2	0.49	0.57	0.27	0.34
cs3	0.67	0.71	0.48	0.52
cs4	0.27	0.42	0.08	0.17

Table 74

TAGG-P Factor Loadings and R^2 values by Grade

Item	Factor Loading					R^2 Value				
	9th	10th	11th	12th	18-21	9th	10th	11th	12th	18-21
sl1	0.72	0.72	0.72	0.71	0.72	0.51	0.52	0.52	0.50	0.52
sl2	0.76	0.78	0.69	0.75	0.78	0.58	0.58	0.52	0.59	0.59
sl3	0.80	0.73	0.76	0.79	0.87	0.64	0.60	0.62	0.64	0.68
sl4	0.76	0.73	0.74	0.76	0.82	0.63	0.56	0.57	0.63	0.61
da1	0.69	0.70	0.68	0.68	0.69	0.48	0.48	0.46	0.46	0.48
da2	0.82	0.91	0.77	0.82	0.93	0.64	0.72	0.64	0.69	0.73
da3	0.62	0.76	0.55	0.63	0.53	0.48	0.49	0.33	0.36	0.31
da4	0.56	0.66	0.52	0.55	0.53	0.29	0.39	0.28	0.31	0.29
p1	0.81	0.80	0.80	0.78	0.79	0.65	0.63	0.63	0.61	0.63
p2	0.91	0.93	0.92	0.91	0.94	0.83	0.83	0.83	0.81	0.83
p3	0.89	0.82	0.88	0.89	0.99	0.80	0.77	0.80	0.80	0.83
p4	0.91	0.88	0.91	0.91	0.94	0.83	0.81	0.83	0.83	0.84
p5	0.82	0.83	0.81	0.82	0.74	0.70	0.68	0.67	0.68	0.63
io1	0.79	0.79	0.79	0.79	0.78	0.62	0.62	0.62	0.82	0.62
io2	0.55	0.61	0.52	0.54	0.55	0.32	0.35	0.28	0.32	0.30
io3	0.79	0.82	0.75	0.79	0.94	0.66	0.63	0.59	0.61	0.70
ga1	0.79	0.77	0.79	0.78	0.79	0.62	0.60	0.62	0.61	0.63
ga2	0.91	0.91	0.88	0.80	0.90	0.79	0.81	0.80	0.82	0.81
ga3	0.95	0.96	0.90	0.92	0.84	0.85	0.88	0.87	0.88	0.85
ga4	0.91	0.88	0.88	0.89	0.89	0.78	0.80	0.81	0.82	0.91

Table 74

TAGG-P Factor Loadings and R^2 values by Grade, continued

	Factor Loading					R^2 Value				
Item	9th	10th	11th	12th	18-21	9th	10th	11th	12th	18-21
ga5	0.87	0.91	0.85	0.86	0.85	0.72	0.77	0.75	0.75	0.74
ga6	0.33	0.44	0.32	0.34	0.30	0.23	0.18	0.10	0.06	0.09
e1	0.81	0.80	0.78	0.81	0.81	0.66	0.65	0.61	0.65	0.65
e2	0.89	0.95	0.88	0.90	0.90	0.80	0.82	0.79	0.78	0.79
e3	0.27	0.34	0.20	0.26	0.15	0.18	0.11	0.04	0.04	0.03
e4	0.42	0.14	0.33	0.39	0.37	0.07	0.02	0.11	0.16	0.14
iep1	0.70	0.68	0.69	0.68	0.70	0.48	0.47	0.47	0.47	0.49
iep2	0.92	0.83	0.93	0.90	0.91	0.78	0.80	0.83	0.84	0.84
iep3	0.92	0.78	0.95	0.90	0.86	0.75	0.77	0.85	0.82	0.80
iep4	0.63	0.46	0.64	0.59	0.62	0.17	0.26	0.40	0.41	0.37
cs1	0.73	0.71	0.72	0.72	0.71	0.53	0.51	0.51	0.52	0.51
cs2	0.61	0.75	0.56	0.62	0.62	0.40	0.48	0.32	0.39	0.38
cs3	0.74	0.75	0.75	0.72	0.80	0.47	0.55	0.56	0.55	0.58
cs4	0.38	0.33	0.32	0.37	0.40	0.03	0.11	0.10	0.15	0.16

Table 75

TAGG-P Factor Loadings and R^2 values by Disability Category

	Factor Loading					R^2 Value				
Item	AUT	ED	ID	OHI	SLD	AUT	ED	ID	OHI	SLD
sl1	0.72	0.72	0.72	0.72	0.72	0.52	0.52	0.52	0.52	0.51
sl2	0.75	0.82	0.78	0.74	0.74	0.64	0.61	0.59	0.55	0.56
sl3	0.80	0.80	0.78	0.80	0.78	0.69	0.65	0.62	0.64	0.63
sl4	0.76	0.80	0.75	0.73	0.75	0.72	0.61	0.57	0.56	0.57
da1	0.69	0.70	0.69	0.69	0.69	0.48	0.49	0.48	0.47	0.47
da2	0.82	0.77	0.91	0.79	0.80	0.69	0.64	0.72	0.66	0.66
da3	0.62	0.65	0.48	0.60	0.60	0.47	0.41	0.27	0.37	0.37
da4	0.56	0.54	0.64	0.61	0.54	0.41	0.60	0.38	0.35	0.30
p1	0.79	0.80	0.79	0.79	0.79	0.63	0.64	0.63	0.63	0.63
p2	0.90	0.94	0.95	0.90	0.89	0.84	0.84	0.83	0.82	0.81
p3	0.89	0.87	0.97	0.89	0.89	0.80	0.79	0.82	0.80	0.79
p4	0.91	0.90	0.99	0.87	0.90	0.86	0.83	0.85	0.81	0.83
p5	0.82	0.83	0.75	0.82	0.82	0.76	0.69	0.63	0.68	0.67
io1	0.77	0.79	0.78	0.78	0.79	0.60	0.62	0.61	0.62	0.63
io2	0.55	0.43	0.34	0.52	0.62	0.05	0.21	0.14	0.28	0.36
io3	0.78	0.81	0.69	0.82	0.75	0.69	0.63	0.54	0.64	0.59
ga1	0.79	0.79	0.78	0.79	0.78	0.62	0.62	0.61	0.62	0.61
ga2	0.90	0.82	0.96	0.88	0.89	0.81	0.78	0.84	0.80	0.81
ga3	0.84	0.84	0.94	0.87	0.95	0.87	0.85	0.88	0.85	0.88
ga4	0.90	0.83	0.84	0.91	0.91	0.80	0.79	0.79	0.82	0.82

Table 75

TAGG-P Factor Loadings and R^2 values by Disability Category, continued

	Factor Loading					R^2 Value				
Item	AUT	ED	ID	OHI	SLD	AUT	ED	ID	OHI	SLD
ga5	0.87	0.81	0.80	0.90	0.87	0.75	0.73	0.81	0.77	0.75
ga6	0.35	0.52	0.33	0.38	0.31	0.09	0.24	0.11	0.14	0.10
e1	0.80	0.81	0.81	0.81	0.80	0.64	0.66	0.66	0.66	0.65
e2	0.89	0.86	0.86	0.87	0.88	0.79	0.77	0.81	0.78	0.79
e3	0.27	0.32	0.30	0.24	0.27	0.80	0.10	0.09	0.06	0.07
e4	0.41	0.32	0.40	0.31	0.44	0.14	0.11	0.16	0.10	0.19
iep1	0.70	0.70	0.69	0.69	0.69	0.49	0.48	0.47	0.48	0.48
iep2	0.64	0.78	0.90	0.90	0.94	0.83	0.78	0.83	0.83	0.84
iep3	0.91	0.80	0.89	0.90	0.95	0.80	0.79	0.82	0.82	0.84
iep4	0.62	0.43	0.52	0.54	0.67	0.44	0.23	0.60	0.31	0.42
cs1	0.72	0.72	0.72	0.72	0.72	0.52	0.52	0.52	0.52	0.52
cs2	0.61	0.71	0.57	0.61	0.61	0.22	0.45	0.34	0.38	0.37
cs3	0.73	0.66	0.73	0.72	0.73	0.53	0.48	0.53	0.52	0.54
cs4	0.37	0.43	0.32	0.35	0.39	0.18	0.17	0.11	0.13	0.15

Table 76

TAGG-F Factor Loadings and R^2 values by Percent of Time in General Education

	Factor Loading							R^2 Value						
Item	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7-100	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7-100
sl1	0.81	0.80	0.81	0.81	0.82	0.81	0.81	0.65	0.65	0.65	0.66	0.67	0.66	0.65
sl2	0.66	0.66	0.66	0.66	0.67	0.66	0.67	0.44	0.43	0.43	0.44	0.45	0.41	0.44
sl3	0.83	0.83	0.83	0.83	0.82	0.83	0.81	0.68	0.68	0.67	0.68	0.67	0.68	0.66
sl4	0.48	0.48	0.48	0.48	0.50	0.49	0.49	0.23	0.23	0.24	0.23	0.25	0.24	0.24
da1	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.70	0.70	0.70	0.70	0.70	0.70	0.70
da2	0.49	0.85	0.86	0.86	0.85	0.85	0.85	0.74	0.73	0.73	0.73	0.72	0.73	0.73
da3	0.54	0.54	0.55	0.55	0.54	0.54	0.55	0.30	0.30	0.30	0.30	0.30	0.30	0.30
da4	0.53	0.56	0.55	0.55	0.46	0.56	0.56	0.31	0.31	0.31	0.31	0.31	0.31	0.31
p1	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.41	0.41	0.41	0.41	0.41	0.41	0.41
p2	0.86	0.86	0.86	0.86	0.86	0.86	0.89	0.75	0.74	0.74	0.74	0.74	0.74	0.74
p3	0.88	0.89	0.88	0.88	0.88	0.88	0.88	0.78	0.78	0.78	0.78	0.78	0.78	0.78
p4	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.74	0.74	0.74	0.74	0.74	0.74	0.74
p5	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.60	0.59	0.59	0.60	0.60	0.59	0.60
io1	0.64	0.65	0.65	0.65	0.65	0.65	0.64	0.41	0.42	0.42	0.42	0.42	0.43	0.42
io2	0.42	0.42	0.43	0.43	0.41	0.42	0.42	0.17	0.18	0.18	0.18	0.17	0.18	0.18
io3	0.61	0.62	0.62	0.62	0.62	0.62	0.62	0.37	0.38	0.38	0.38	0.38	0.38	0.38
ga1	0.76	0.77	0.77	0.77	0.77	0.77	0.76	0.58	0.59	0.58	0.59	0.59	0.59	0.58
ga2	0.79	0.79	0.79	0.79	0.79	0.80	0.79	0.63	0.63	0.63	0.63	0.63	0.64	0.62
ga3	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.69	0.69	0.69	0.69	0.39	0.69	0.69
ga4	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.60	0.60	0.60	0.60	0.60	0.60	0.61

Table 76

TAGG-F Factor Loadings and R^2 values by Percent of Time in General Education, continued

	Factor Loadings							R^2 Values						
Item	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7-100	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7-100
ga5	0.86	0.86	0.85	0.85	0.85	0.86	0.85	0.73	0.73	0.73	0.73	0.73	0.73	0.73
ga6	0.26	0.26	0.27	0.27	0.27	0.27	0.26	0.07	0.07	0.07	0.07	0.07	0.07	0.07
e1	0.73	0.73	0.73	0.74	0.73	0.74	0.73	0.54	0.54	0.54	0.54	0.54	0.54	0.54
e2	0.89	0.88	0.88	0.88	0.88	0.87	0.88	0.76	0.77	0.77	0.77	0.78	0.76	0.77
e3	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.03	0.03	0.03	0.03	0.03	0.03	0.03
e4	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.05	0.05	0.05	0.05	0.05	0.05	0.05
iep1	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.52	0.52	0.52	0.52	0.52	0.52	0.52
iep2	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.75	0.75	0.76	0.76	0.75	0.76	0.75
iep3	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.77	0.77	0.78	0.78	0.77	0.77	0.77
iep4	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.60	0.30	0.31	0.90	0.30	0.30	0.31
cs1	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.37	0.37	0.37	0.37	0.37	0.37	0.37
cs2	0.56	0.55	0.55	0.55	0.55	0.55	0.56	0.31	0.30	0.30	0.30	0.30	0.31	0.31
cs3	0.65	0.64	0.65	0.65	0.65	0.65	0.65	0.42	0.42	0.42	0.42	0.42	0.42	0.42
cs4	0.38	0.39	0.39	0.39	0.39	0.39	0.38	0.14	0.15	0.15	0.15	0.15	0.15	0.15

Table 77

TAGG-F Factor Loadings and R^2 values by Transition Education

	Factor Loading		R^2 Value	
Item	Little to no Transition Education	Some Transition Education	Little to no Transition Education	Some Transition Education
sl1	0.81	0.81	0.66	0.65
sl2	0.66	0.66	0.44	0.44
sl3	0.82	0.81	0.67	0.66
sl4	0.49	0.50	0.24	0.25
da1	0.84	0.84	0.70	0.70
da2	0.85	0.86	0.76	0.73
da3	0.55	0.55	0.30	0.30
da4	0.56	0.56	0.31	0.31
p1	0.64	0.64	0.41	0.41
p2	0.86	0.86	0.74	0.74
p3	0.88	0.88	0.78	0.78
p4	0.86	0.86	0.74	0.74
p5	0.77	0.77	0.60	0.60
io1	0.65	0.64	0.42	0.41
io2	0.42	0.42	0.18	0.17
io3	0.62	0.62	0.38	0.38
ga1	0.77	0.76	0.59	0.58
ga2	0.79	0.79	0.63	0.63
ga3	0.83	0.83	0.69	0.70
ga4	0.78	0.78	0.60	0.61

Table 77

TAGG-F Factor Loadings and R^2 values by Transition Education, continued

	Factor Loading		R^2 Value	
Item	Little to no Transition Education	Some Transition Education	Little to no Transition Education	Some Transition Education
ga5	0.85	0.85	0.73	0.73
ga6	0.27	0.26	0.07	0.07
e1	0.74	0.74	0.54	0.54
e2	0.88	0.88	0.77	0.77
e3	0.16	0.17	0.03	0.03
e4	0.23	0.23	0.05	0.05
iep1	0.72	0.72	0.52	0.52
iep2	0.87	0.87	0.75	0.75
iep3	0.88	0.88	0.77	0.77
iep4	0.55	0.55	0.30	0.60
cs1	0.61	0.61	0.37	0.37
cs2	0.55	0.55	0.30	0.31
cs3	0.65	0.65	0.42	0.42
cs4	0.39	0.39	0.15	0.15

Table 78

TAGG-F Factor Loadings and R^2 values by Grade

Item	Factor Loading				R^2 Value			
	9th	10th	11th	12th	9th	10th	11th	12th
sl1	0.82	0.81	0.81	0.81	0.67	0.65	0.65	0.66
sl2	0.66	0.67	0.68	0.65	0.43	0.45	0.43	0.42
sl3	0.82	0.82	0.80	0.83	0.68	0.67	0.64	0.68
sl4	0.49	0.49	0.51	0.48	0.24	0.24	0.26	0.23
da1	0.84	0.84	0.84	0.84	0.70	0.70	0.70	0.70
da2	0.85	0.85	0.85	0.85	0.72	0.76	0.73	0.73
da3	0.54	0.54	0.54	0.55	0.29	0.30	0.29	0.30
da4	0.55	0.56	0.56	0.55	0.30	0.32	0.31	0.31
p1	0.63	0.64	0.64	0.64	0.40	0.41	0.41	0.40
p2	0.88	0.86	0.86	0.86	0.74	0.74	0.74	0.74
p3	0.86	0.88	0.88	0.88	0.77	0.78	0.78	0.78
p4	0.86	0.86	0.86	0.86	0.74	0.74	0.74	0.74
p5	0.77	0.77	0.77	0.77	0.59	0.59	0.59	0.59
io1	0.65	0.65	0.65	0.65	0.42	0.42	0.42	0.42
io2	0.43	0.42	0.43	0.42	0.19	0.18	0.18	0.18
io3	0.62	0.62	0.62	0.61	0.38	0.38	0.38	0.37
ga1	0.77	0.76	0.76	0.77	0.59	0.58	0.58	0.59
ga2	0.79	0.79	0.79	0.79	0.63	0.63	0.63	0.62
ga3	0.83	0.83	0.83	0.83	0.69	0.67	0.69	0.69
ga4	0.77	0.78	0.78	0.77	0.60	0.70	0.60	0.60

Table 78

TAGG-F Factor Loadings and R^2 values by Grade, continued

	Factor Loading				R^2 Value			
Item	9th	10th	11th	12th	9th	10th	11th	12th
ga5	0.85	0.85	0.85	0.85	0.72	0.73	0.73	0.73
ga6	0.24	0.27	0.27	0.25	0.06	0.07	0.07	0.06
e1	0.73	0.74	0.74	0.74	0.54	0.54	0.54	0.54
e2	0.88	0.88	0.88	0.87	0.77	0.77	0.77	0.76
e3	0.17	0.17	0.17	0.16	0.03	0.03	0.03	0.03
e4	0.22	0.23	0.23	0.21	0.05	0.05	0.05	0.05
iep1	0.72	0.72	0.72	0.72	0.52	0.52	0.52	0.52
iep2	0.87	0.87	0.87	0.87	0.07	0.75	0.75	0.75
iep3	0.88	0.88	0.88	0.88	0.77	0.78	0.77	0.77
iep4	0.54	0.55	0.55	0.54	0.29	0.30	0.31	0.60
cs1	0.62	0.61	0.61	0.61	0.38	0.37	0.37	0.37
cs2	0.55	0.55	0.55	0.55	0.30	0.60	0.30	0.30
cs3	0.64	0.65	0.65	0.65	0.41	0.42	0.42	0.42
cs4	0.38	0.39	0.39	0.37	0.14	0.15	0.15	0.14

Table 79

TAGG-F Factor Loadings and R^2 values by Disability Category

	Factor Loading			R^2 Value		
Item	ID	OHI	SLD	ID	OHI	SLD
sl1	0.80	0.83	0.81	0.63	0.68	0.66
sl2	0.67	0.67	0.67	0.44	0.45	0.45
sl3	0.81	0.82	0.79	0.66	0.67	0.43
sl4	0.49	0.50	0.48	0.24	0.25	0.23
da1	0.84	0.84	0.83	0.70	0.70	0.70
da2	0.85	0.85	0.85	0.73	0.73	0.73
da3	0.55	0.55	0.55	0.30	0.30	0.31
da4	0.56	0.56	0.55	0.32	0.31	0.30
p1	0.64	0.64	0.64	0.41	0.41	0.39
p2	0.89	0.86	0.86	0.75	0.75	0.74
p3	0.88	0.88	0.88	0.78	0.78	0.77
p4	0.86	0.86	0.85	0.74	0.74	0.73
p5	0.77	0.77	0.76	0.59	0.59	0.57
io1	0.65	0.65	0.64	0.42	0.42	0.41
io2	0.42	0.42	0.42	0.18	0.18	0.18
io3	0.62	0.62	0.60	0.38	0.38	0.36
ga1	0.77	0.77	0.75	0.59	0.59	0.56
ga2	0.79	0.79	0.78	0.63	0.63	0.61
ga3	0.83	0.83	0.82	0.69	0.69	0.68
ga4	0.77	0.78	0.76	0.60	0.60	0.58

Table 79

TAGG-F Factor Loadings and R^2 values by Disability Category, continued

	Factor Loading			R^2 Value		
Item	ID	OHI	SLD	ID	OHI	SLD
ga5	0.86	0.86	0.85	0.73	0.73	0.72
ga6	0.27	0.27	0.27	0.07	0.07	0.07
e1	0.73	0.74	0.72	0.54	0.54	0.52
e2	0.88	0.88	0.88	0.77	0.77	0.77
e3	0.16	0.16	0.15	0.03	0.03	0.02
e4	0.22	0.23	0.19	0.05	0.05	0.04
iep1	0.72	0.72	0.71	0.52	0.52	0.51
iep2	0.87	0.87	0.87	0.73	0.75	0.76
iep3	0.88	0.88	0.87	0.78	0.77	0.76
iep4	0.55	0.55	0.54	0.60	0.31	0.29
cs1	0.61	0.61	0.58	0.37	0.38	0.34
cs2	0.55	0.55	0.56	0.30	0.30	0.32
cs3	0.64	0.65	0.64	0.41	0.42	0.41
cs4	0.39	0.39	0.37	0.15	0.15	0.14

Table 80

TAGG-S Factor Loadings and R^2 values by Percent of Time in General Education

	Factor Loading							R^2 Value						
Item	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100
SCSL1	0.54	0.55	0.57	0.53	0.53	0.50	0.52	0.35	0.36	0.38	0.34	0.36	0.35	0.35
SCSL2	0.56	0.37	0.53	0.42	0.63	0.30	0.38	0.32	0.20	0.28	0.19	0.14	0.14	0.16
SCSL3	0.39	0.40	0.54	0.44	0.48	0.51	0.51	0.17	0.19	0.24	0.19	0.18	0.16	0.21
SCSL4	0.42	0.21	0.51	0.31	0.36	0.30	0.27	0.01	0.04	0.04	0.05	0.01	0.03	0.05
SCSL5	0.32	0.59	0.49	0.60	0.57	0.60	0.56	0.49	0.41	0.41	0.40	0.36	0.32	0.36
SCSL6	0.36	0.51	0.38	0.42	0.50	0.57	0.51	0.41	0.18	0.32	0.42	0.22	0.30	0.29
SCSL7	0.50	0.46	0.33	0.60	0.51	0.49	0.44	0.34	0.08	0.36	0.15	0.32	0.32	0.27
SCSL8	0.22	0.44	0.21	0.60	0.24	0.34	0.31	0.09	0.08	0.05	0.00	0.01	0.00	0.00
da1	0.62	0.61	0.60	0.61	0.63	0.61	0.61	0.41	0.42	0.43	0.42	0.43	0.42	0.41
da2	0.77	0.78	0.83	0.81	0.76	0.78	0.75	0.39	0.39	0.42	0.38	0.42	0.47	0.45
da3	0.47	0.65	0.79	0.76	0.69	0.74	0.72	0.25	0.23	0.26	0.10	0.27	0.23	0.23
da4	0.59	0.55	0.79	0.72	0.67	0.74	0.73	0.16	0.23	0.15	0.14	0.11	0.19	0.16
p1	0.63	0.63	0.65	0.63	0.64	0.64	0.64	0.39	0.38	0.37	0.38	0.41	0.39	0.39
p2	0.45	0.35	0.53	0.36	0.43	0.43	0.42	0.60	0.59	0.62	0.64	0.57	0.61	0.58
p3	0.19	0.06	0.17	0.17	0.24	0.36	0.29	0.60	0.44	0.53	0.55	0.46	0.50	0.49
p4	0.36	0.35	0.72	0.55	0.56	0.59	0.61	0.40	0.35	0.52	0.51	0.45	0.51	0.49
p5	0.54	0.58	0.56	0.46	0.54	0.65	0.56	0.34	0.28	0.50	0.42	0.38	0.40	0.39
io1	0.57	0.59	0.56	0.57	0.57	0.58	0.57	0.26	0.26	0.27	0.25	0.27	0.27	0.26
io2	0.59	0.81	0.69	0.77	0.70	0.67	0.69	0.04	0.01	0.02	0.05	0.09	0.17	0.10
io3	0.58	0.43	0.54	0.54	0.49	0.48	0.52	0.30	0.22	0.53	0.38	0.43	0.43	0.43

Table 80

TAGG-S Factor Loadings and R^2 values by Percent of Time in General Education, continued

	Factor Loadings							R^2 Values						
Item	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100	<24.9	25.0 - 39.2	39.3 - 53.6	53.4 - 67.9	68.0 - 82.3	82.4 - 96.6	96.7- 100
ga1	0.49	0.49	0.50	0.48	0.46	0.48	0.48	0.34	0.33	0.35	0.33	0.34	0.35	0.34
ga2	0.23	0.25	-0.05	0.00	0.09	0.17	0.01	0.36	0.38	0.14	0.35	0.27	0.24	0.23
ga3	0.67	0.76	0.82	0.66	0.91	0.75	0.78	0.41	0.54	0.42	0.51	0.45	0.39	0.43
ga4	0.86	0.86	0.92	0.81	0.96	0.82	0.84	0.30	0.18	0.24	0.28	0.23	0.19	0.23
ga5	0.06	0.20	0.07	0.06	0.04	0.06	0.07	0.52	0.44	0.55	0.49	0.49	0.45	0.47
ga6	0.18	0.13	0.23	0.01	0.13	0.09	0.06	0.01	0.02	0.00	0.01	0.00	0.01	0.02
e1	0.68	0.68	0.69	0.68	0.67	0.69	0.68	0.62	0.63	0.62	0.63	0.65	0.64	0.64
e2	0.75	0.67	0.78	0.73	0.85	0.82	0.81	0.81	0.79	0.81	0.80	0.80	0.78	0.78
e3	0.77	0.75	0.72	0.77	0.84	0.83	0.79	0.04	0.04	0.00	0.00	0.00	0.00	0.00
e4	0.45	0.36	0.33	0.33	0.48	0.47	0.42	0.02	0.01	0.04	0.00	0.01	0.01	0.00
iep1	0.62	0.62	0.63	0.62	0.63	0.62	0.62	0.46	0.46	0.48	0.46	0.45	0.48	0.49
iep2	0.63	0.41	0.62	0.63	0.54	0.57	0.55	0.60	0.54	0.62	0.58	0.66	0.65	0.63
iep3	0.59	0.33	0.65	0.37	0.32	0.58	0.55	0.63	0.61	0.59	0.62	0.66	0.21	0.64
iep4	0.32	0.31	0.26	0.02	0.09	0.00	0.06	0.22	0.13	0.12	0.12	0.22	0.32	0.18

Table 81

TAGG-S Factor Loadings and R^2 values by Transition Education

	Factor Loading		R^2 Value	
Item	Little to no Transition Education	Some Transition Education	Little to no Transition Education	Some Transition Education
SCSL1	0.52	0.55	0.35	0.37
SCSL2	0.35	0.47	0.01	0.19
SCSL3	0.49	0.52	0.14	0.23
SCSL4	0.24	0.44	0.04	0.08
SCSL5	0.54	0.50	0.43	0.39
SCSL6	0.41	0.48	0.28	0.08
SCSL7	0.31	0.48	0.27	0.27
SCSL8	0.29	0.26	0.05	0.10
da1	0.62	0.61	0.42	0.42
da2	0.81	0.72	0.42	0.43
da3	0.69	0.65	0.13	0.23
da4	0.65	0.69	0.15	0.15
p1	0.66	0.63	0.40	0.38
p2	0.44	0.44	0.63	0.60
p3	0.50	0.25	0.45	0.49
p4	0.63	0.52	0.42	0.47
p5	0.47	0.60	0.48	0.39
io1	0.56	0.58	0.25	0.28
io2	0.64	0.67	0.05	0.09
io3	0.54	0.56	0.45	0.40

Table 81

TAGG-S Factor Loadings and R^2 values by Transition Education, continued

	Factor Loading		R^2 Value	
Item	Little to no Transition Education	Some Transition Education	Little to no Transition Education	Some Transition Education
ga1	0.48	0.51	0.32	0.35
ga2	0.11	0.02	0.18	0.26
ga3	0.81	0.80	0.34	0.44
ga4	0.93	0.81	0.21	0.24
ga5	-0.01	0.11	0.50	0.49
ga6	0.13	0.05	0.06	0.09
e1	0.68	0.68	0.64	0.64
e2	0.84	0.73	0.82	0.78
e3	0.83	0.80	0.00	0.01
e4	0.56	0.30	0.01	0.14
iep1	0.63	0.64	0.42	0.46
iep2	0.53	0.56	0.65	0.63
iep3	0.54	0.59	0.65	0.64
iep4	-0.05	0.12	0.29	0.17

Table 82

TAGG-S Factor Loadings and R^2 values by Grade

	Factor Loading					R^2 Value				
Item	9th	10th	11th	12th	18-21	9th	10th	11th	12th	18-21
SCSL1	0.51	0.55	0.52	0.55	0.55	0.35	0.36	0.34	0.37	0.36
SCSL2	0.22	0.41	0.35	0.43	0.28	0.21	0.16	0.16	0.21	0.10
SCSL3	0.54	0.50	0.52	0.48	0.59	0.09	0.23	0.20	0.23	0.25
SCSL4	0.16	0.21	0.26	0.42	0.19	0.05	0.42	0.04	0.14	0.01
SCSL5	0.60	0.46	0.54	0.56	0.47	0.20	0.35	0.38	0.41	0.31
SCSL6	0.59	0.36	0.49	0.48	0.60	0.15	0.24	0.33	0.30	0.12
SCSL7	0.42	0.30	0.43	0.47	0.41	0.18	0.27	0.26	0.30	0.18
SCSL8	0.37	0.14	0.31	0.29	0.26	0.00	0.02	0.01	0.01	0.01
da1	0.62	0.61	0.61	0.60	0.62	0.43	0.42	0.42	0.42	0.42
da2	0.83	0.76	0.80	0.75	0.76	0.43	0.42	0.42	0.46	0.32
da3	0.76	0.81	0.72	0.66	0.72	0.16	0.25	0.18	0.29	0.25
da4	0.58	0.69	0.71	0.73	0.58	0.18	0.16	0.14	0.16	0.15
p1	0.64	0.63	0.64	0.64	0.64	0.40	0.39	0.39	0.38	0.40
p2	0.52	0.47	0.52	0.39	0.39	0.62	0.59	0.60	0.58	0.58
p3	0.23	0.61	0.30	0.23	0.64	0.50	0.54	0.48	0.44	0.50
p4	0.62	0.46	0.58	0.53	0.61	0.33	0.45	0.48	0.51	0.38
p5	0.63	0.63	0.58	0.53	0.67	0.40	0.36	0.40	0.39	0.39
io1	0.57	0.56	0.57	0.57	0.58	0.26	0.27	0.28	0.24	0.27
io2	0.72	0.76	0.73	0.63	0.77	0.05	0.12	0.10	0.08	0.16
io3	0.64	0.58	0.48	0.54	0.39	0.47	0.35	0.41	0.39	0.53

Table 82

TAGG-S Factor Loadings and R^2 values by Grade, continued

	Factor Loading					R^2 Value				
Item	9th	10th	11th	12th	18-21	9th	10th	11th	12th	18-21
ga1	0.47	0.50	0.46	0.50	0.51	0.35	0.35	0.35	0.32	0.35
ga2	0.07	0.15	0.13	0.00	0.19	0.20	0.21	0.27	0.27	0.33
ga3	0.80	0.70	0.85	0.76	0.84	0.38	0.45	0.49	0.38	0.48
ga4	0.82	0.87	0.92	0.80	0.92	0.28	0.60	0.19	0.25	0.13
ga5	0.61	0.07	-0.02	0.11	0.13	0.52	0.46	0.49	0.49	0.53
ga6	0.37	0.20	0.07	0.12	-0.02	0.10	0.13	0.05	0.01	0.08
e1	0.86	0.69	0.68	0.67	0.69	0.39	0.62	0.83	0.69	0.63
e2	0.80	0.75	0.81	0.78	0.69	0.78	0.81	0.63	0.78	0.82
e3	0.92	0.81	0.79	0.78	0.71	0.09	0.00	0.00	0.01	0.02
e4	0.36	0.34	0.41	0.45	0.24	0.11	0.38	0.00	0.01	0.00
iep1	0.61	0.63	0.62	0.63	0.62	0.48	0.47	0.46	0.45	0.47
iep2	0.40	0.51	0.62	0.57	0.42	0.62	0.60	0.64	0.61	0.51
iep3	0.48	0.54	0.52	0.57	0.41	0.70	0.13	0.63	0.63	0.53
iep4	-0.03	0.15	0.09	0.13	0.06	0.13	0.35	0.18	0.20	0.06

Table 83

TAGG-S Factor Loadings and R^2 values by Disability Category

	Factor Loading					R^2 Value				
Item	AUT	ED	ID	OHI	SLD	AUT	ED	ID	OHI	SLD
SCSL1	0.55	0.54	0.56	0.53	0.49	0.37	0.35	0.37	0.35	0.34
SCSL2	0.40	0.49	0.13	0.39	0.37	0.22	0.19	0.03	0.20	0.16
SCSL3	0.62	0.32	0.49	0.57	0.50	0.31	0.14	0.26	0.24	0.21
SCSL4	0.32	0.54	0.28	0.37	0.27	0.12	0.23	0.08	0.10	0.05
SCSL5	0.48	0.59	0.46	0.67	0.58	0.42	0.35	0.44	0.21	0.36
SCSL6	0.31	0.44	0.67	0.51	0.53	0.11	0.32	0.35	0.23	0.29
SCSL7	0.44	0.29	0.40	0.59	0.45	0.16	0.25	0.24	0.19	0.27
SCSL8	0.26	0.23	0.24	0.41	0.31	0.09	0.00	0.09	0.06	0.00
da1	0.62	0.62	0.61	0.61	0.61	0.42	0.43	0.42	0.46	0.40
da2	0.75	0.84	0.84	0.80	0.76	0.34	0.38	0.38	0.45	0.46
da3	0.66	0.69	0.70	0.69	0.82	0.22	0.12	0.22	0.27	0.23
da4	0.41	0.71	0.66	0.79	0.83	0.12	0.12	0.12	0.18	0.16
p1	0.64	0.65	0.63	0.62	0.64	0.40	0.39	0.38	0.38	0.39
p2	0.57	0.45	0.38	0.46	0.42	0.58	0.63	0.64	0.60	0.58
p3	0.14	0.09	0.26	0.22	0.29	0.41	0.47	0.48	0.45	0.49
p4	0.53	0.51	0.49	0.43	0.61	0.19	0.49	0.45	0.52	0.50
p5	0.61	0.54	0.63	0.59	0.56	0.41	0.46	0.36	0.22	0.39
io1	0.58	0.47	0.57	0.57	0.57	0.27	0.26	0.26	0.27	0.25
io2	0.74	0.50	0.83	0.76	0.69	0.07	0.01	0.07	0.08	0.10
io3	0.43	0.55	0.53	0.46	0.52	0.41	0.39	0.45	0.28	0.43

Table 83

TAGG-S Factor Loadings and R^2 values by Disability Category, continued

	Factor Loading					R^2 Value				
Item	AUT	ED	ID	OHI	SLD	AUT	ED	ID	OHI	SLD
ga1	0.49	0.51	0.49	0.58	0.48	0.34	0.34	0.35	0.34	0.33
ga2	0.40	0.08	0.24	0.28	0.00	0.30	0.23	0.29	0.60	0.26
ga3	0.98	0.76	0.78	0.66	0.80	0.47	0.26	0.54	0.48	0.44
ga4	0.99	0.81	0.98	0.80	0.85	0.16	0.23	0.28	0.21	0.23
ga5	0.18	0.05	0.08	0.02	0.07	0.47	0.51	0.56	0.41	0.47
ga6	0.08	0.16	0.13	0.14	0.06	0.30	0.00	0.15	0.76	0.02
e1	0.69	0.69	0.68	0.69	0.68	0.64	0.62	0.60	0.60	0.66
e2	0.77	0.82	0.72	0.84	0.80	0.82	0.80	0.84	0.68	0.78
e3	0.83	0.80	0.79	0.83	0.79	0.02	0.00	0.01	0.00	0.00
e4	0.42	0.27	0.42	0.52	0.42	0.01	0.04	0.01	0.02	0.00
iep1	0.62	0.62	0.62	0.60	0.62	0.48	0.47	0.46	0.47	0.45
iep2	0.38	0.61	0.62	0.50	0.53	0.60	0.63	0.57	0.65	0.63
iep3	0.36	0.52	0.50	0.41	0.55	0.65	0.63	0.64	0.65	0.64
iep4	0.36	-0.10	0.34	0.30	0.06	0.17	0.08	0.16	0.25	0.18

***TAGG-P* factor loadings and R^2 values.** The factor loadings and R^2 values for the three *TAGG* versions indicated most of the *TAGG* items represent the constructs and make meaningful contributions to a test taker's *TAGG* scores. However, a few of the *TAGG* items did not meet the criteria, meaning they were not good indicators of the constructs the *TAGG* purports to measure or did not make meaningful contributions to *TAGG* scores. On the *TAGG-P*, an item asking if the student had attained an annual transition goal had factor loadings ranging from .27-.56 and was below .35 for 13 of the 19 groups I examined for this study. A second item that addressed unpaid employment had factor loadings ranging from .10-.35 and was less than .35 for 18 of the 19 groups I examined for this study. Neither of these items had R^2 values that were less than .15, but relative to the R^2 values of other items on the *TAGG-P* (which were generally greater than .50), these two items did not appear to make as great a contribution to *TAGG* scores.

***TAGG-F* factor loadings and R^2 values.** On the *TAGG-F*, three items did not exhibit acceptable levels of factor loadings and R^2 values for any of the groups I examined for this study. One Goal Attainment construct item (attaining an annual transition goal) and two Employment construct items (having unpaid employment and having paid employment) had both below acceptable factor loadings and R^2 values for all of the groups. About half of the families who participated in this study did not respond to the item about whether their students had attained an annual transition goal (Table 58) suggesting that families do not understand the item—either they may not know what an annual transition goal is, or they do not know if their child has attained any of his annual transition goals.

The two Employment construct items that did not meet the criteria for inclusion on the *TAGG-F* present valuable family perspective. When the *TAGG* was developed, any research-based behaviors that were associated with postsecondary education and employment were included on the original 75-item *TAGG* assessment. The behavior for the final 34-item *TAGG* were based on professional educators perceptions, and it appears those professional educators viewed having an unpaid job and/or a paid job as valuable behaviors for measuring their students' preparedness for postsecondary education and competitive employment.

In contrast to the educators who participated in *TAGG* development, families who participated do not appear to view having a paid or unpaid job as related to future employment or postsecondary education. This may be related to their perceptions of their students' difficulty with school. For the families of students with disabilities, school is already more challenging than for students who do not have disabilities. It could be that families view having paid (or unpaid) employment during high school as an unnecessary burden on the student that interferes with their academic success.

The *TAGG* developers should use caution when generating these two goals on the *TAGG-F* scoring profile. First, it does not appear, from this study, that families endorse these behaviors, and including them as goals suggested by their scoring profile is a misrepresentation of their perspectives. Second, having a job during high school requires students and families to put forth substantial effort, and if having a job does not contribute to improvement in a student's Employment construct ability, working toward this goal is not a good use of the student and his family's time.

***TAGG-S* factor loadings and R^2 values.** On the *TAGG-S*, four items did not exhibit acceptable levels of factor loadings and R^2 values for most of the groups I examined for this study. Three Goal Attainment construct items (breaking long-term goals into short-term goals, moving on to the next goal after one goal was met, and attaining at least one annual transition goal) and one Involvement in the IEP item (leading the IEP meeting) had both below acceptable factor loadings and R^2 values for nearly all of the groups I examined. There is evidence that students with disabilities do not have the skills associated with the goal attainment process without instruction (Martin, Martin, & Osmani, 2013), so it is likely that students do not know that breaking long-term goals into short-term goals and moving onto the next goals are part of the goal attainment process.

With regards to students' perception that leading their IEP meeting is not a good indicator of their Involvement in the IEP, it is possible that students rely on educators to lead the meeting because of the infrequency with which they are held and the complicated paperwork involved. Leading the meeting may be an action students who participated in this study associated with professional educators or they may not make the same meaning of "lead" as the *TAGG* developers.

Factor loading and R^2 value implications. Like any other transition assessment, the *TAGG* is primarily concerned with identifying behaviors students cannot do so that goals may be generated from those behaviors. If the purpose of the *TAGG* were just to measure how much a student could do, retaining items that do not meet the factor loading and R^2 value criteria simply means having educators, families, and students respond to a few extra items. But in the case of transition assessments, including the *TAGG*, the items are intended to become goals that students will work on for the next school year.

Including items on a transition assessment that do not exhibit a meaningful contribution to improving a student's ability within the constructs risks having the student spend the next year learning and practicing a skill that is not actually going to make a difference in his latent ability. Additionally, the student's educators and family will have used their time and resources supporting the student to learn a behavior that is not meaningfully related to the trait they hope to improve.

For example, educators who participated in this study did not appear to see the behavior of attaining a transition goal as particularly meaningful to improving a students' Goal Attainment ability, therefore, does it make sense for the student to spend the next school year learning to attain his annual transition goals, or if the educator believes Goal Attainment is an area a student needs to improve, are there other behaviors the student needs to learn? Attaining an annual transition goal is the conceptually most difficult item in the Goal Attainment construct, and if a scoring profile suggests this behavior as a goal, it is assumed that the other behaviors included in the Goal Attainment construct have been met (as described in appendix A). From the perspective of this researcher, it seems like students would be best served working on the behaviors that will most improve their ability, which is not necessarily the most difficult to learn behavior. Now that the *TAGG* developers have ample data for understanding the contribution of the *TAGG* behaviors to the constructs, they may want to consider suggesting the goal behaviors that will have the most impact on the areas in which educators, families, and students would like to see improvement.

Future Research and Development

The *TAGG* developers also need to consider the order in which the items are suggested as goals. In its current form, the *TAGG* suggests goals based on their difficulty. Now that data exists that can be analyzed to understand the contribution of the *TAGG* behaviors to construct scores, it may be useful to consider suggesting goals based on how much impact they have on construct scores. For example, accessing social services such as vocational rehabilitation is the most difficult item in its construct, but on all three versions of the *TAGG*, it seems to make little contribution to its construct score, especially for students who spend most of their time in general education and likely have a low level of need. Simply stated, working on this goal is not a good use of time for students who are likely to enter postsecondary education and competitive employment.

The *TAGG* developers wish to use the *TAGG* data set as a snapshot of how students with disabilities are performing on the nonacademic behaviors associated with postsecondary employment and further education. The developers now have evidence that the *TAGG* adequately measures students across several disabilities, grades, and levels of participation in general education. However, before scores can be compared for each of the groups, one more step must be taken that will ensure that having a score of zero on a construct indicates the student indeed has zero ability. To do this, the models must be further constrained and tested. Until this is done, the developers should at most withhold making mean comparisons, or at a minimum, acknowledge that the *TAGG* has not be tested for equality of the scale (i.e., scalar invariance) across the groups being compared.

Conclusion

The development of an assessment means developing a framework that attempts to give structure to behaviors the developers believe to be related. For example, the *TAGG* developers sought to tie together the behaviors they believe are associated with postsecondary employment and further education for students with disabilities. As this study has shown, establishing a framework for an assessment extends far beyond selecting the items and assessing their quantitative relation with the framework's constructs.

Attention must now be paid to whom the *TAGG* is designed for, how the *TAGG* will be used, and whose voice should guide development of the *TAGG*. The *TAGG* developers did a strong job selecting the items to be used on the assessment, but must now adjust their focus for suggesting goals on the scoring profile. Specifically, while educators might be the most knowledgeable about transition skills, families and students have the greatest stake in the activities the *TAGG* profile will suggest be included on the IEP. Actively including those voices extends beyond assessing them and into utilization of their test scores.

This study allowed for a deeper understanding of the meaning of the *TAGG* framework when the assessment was completed by educators of students who spend more time in general education versus educators whose students spend less time in general education. Likewise, it allowed for an understanding of how students view the *TAGG* behaviors as they spend more or less time in general education or have more or less experience practicing them. It was interesting to find that educators expected the same behaviors from students who spend more time in general education than in special education, but that there were some differences in expectations for students in lower grades than upper grades. Had the effort not been made to develop several groups, this could not have been examined.

The broader goal of this study was to present an approach to assessing measurement invariance that used a variable that was believed to present a strong threat to the validity of the scores produced by the test. The study was guided heavily by an applied approach, and consistently asked not whether the findings were significant, but whether the findings were *meaningful*, and if so, what the meaning of the findings were. The *TAGG* framework is being distributed as a new approach to teaching transition education for a very broad range of students, and therefore deserves inquiry to ensure it is indeed a sound approach to teaching transition education.

The methods of inquiry used in this study may be applied to any theoretical framework, but are generally reserved for providing evidence that an assessment is appropriate when arbitrary threats to validity are selected, such as ethnicity, gender, or socio-economic status. In the future, when time and resources are being dedicated to the level of analysis undertaken in a multigroup confirmatory factor analysis, it serves the education community well to take an applied approach that will provide a meaningful understanding of the assessment being reviewed.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational & Psychological Testing (US). (1999). *Standards for educational and psychological testing*. Madison, WI: American Educational Research Association.

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1954). *Technical recommendations for psychological tests and diagnostic techniques*. Washington, DC.

Anderson-Loeb, P. (1996). *Independent living scales*. San Antonio: Pearson Publishing.

Becker, R. L. (2000). *Reading-Free Vocational Interest Inventory* (2nd ed.). Columbus, OH: Elbern Publishing.

Becker, R.L. (2005) *Becker Work Adjustment Profile: 2.* Cicero, NY: Program Development Associates.

Bennett, R. K. (2006). *Bennett Mechanical Comprehension Test*. San Antonio: Pearson Publishing.

Bennet, G. K., Seashore, H. G., & Wesman, A. G. (1947). *Differential aptitude test*. New York: Psychological Corp.

Berry, W. D. (1993). *Understanding regression assumptions*. Thousand Oaks, CA: Sage.

Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, *21*, 33-46.

Brady, M. P., & Rosenberg, H. (2002). Job observation and behavior scale: A supported employment assessment instrument. *Education and Training in Mental Retardation and Developmental Disabilities*, *37*, 427-33.

Brady, R. P. (2007). *Picture interest career survey*. St. Paul, MN: JIST Works.

Bressani, R. V., & Downs, A. C. (2002). Youth independent living assessment: Testing the equivalence of web and paper/pencil versions of the Ansell–Casey Life Skills Assessment. *Computers in Human Behavior*, *18*, 453-464.

Bruininks, R. H., Hill, B. K., Weatherman, R. F., & Woodcock, R. W. (1986). *Inventory for client and agency planning*. Allen, TX: DLM Teaching Resources.

Bucher, D. E., & Brolin, D. E. (1987). The life-centered career education (LCCE) inventory: A curriculum-based, criterion-related assessment instrument. *Assessment for Effective Intervention*, *12*, 131-141.

Byrne, B.M., Shavelson, R. J., & Methuen, B. (1989) Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456-466.

Carter, E.W., Lane, K.L., Pierson, M., & Glasser, B. (2006). Self-determination skills of transition-age youth with emotional disturbances and learning disabilities. *Exceptional Children, 72*, 333-346.

Carter, E. W., Owens, L., Trainor, A. A., Sun, Y., & Swedeen, B. (2009). Self-determination skills and opportunities of adolescents with severe intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disabilities*, *114*, 179-192. doi:10.1352/1944-7558-114.3.179

Casey Family Program. (2002). *Ansell-Casey Life Skills Assessment Lifeskills Guidebook Manual*. Seattle, WA: Casey Family Programs.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245-276.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, *9*, 233-255.

Clark, G. M., & Patton, J. R. (1997). *Transition planning inventory: Administration and resource guide*. Austin, TX: Pro-Ed.

Cohen, J. (1992a). Statistical power analysis. *Current Directions in Psychological Science*, *1*, 98-101.

Cohen, J. (1992b). A power primer. *Psychological Bulletin*, *112*, 155.

Cohen, A. S., Gregg, N., & Deng, M. (2005). The role of extended time and item content on a high-stakes mathematics test. *Learning Disabilities Research and Practice, 20*, 225-233.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98.

Crewe, N. M., & Athelstan, G. T. (1981). *Personal capacities questionnaire*. Minneapolis, MN: University of Minnesota.

Crites, J. O. (1978). *Career maturity inventory*. Monterey, CA: CTB/McGraw-Hill.

Cronbach, L. J. (1947). Test "reliability": Its meaning and determination. *Psychometrika*, *12*, 1-16.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

de Bildt, A., Kraijer, D., Sytema, S., & Minderaa, R. (2005). The psychometric properties of the Vineland Adaptive Behavior Scales in children and adolescents with mental retardation. *Journal of Autism and Developmental disorders*, *35*, 53-62.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*, 121-149. doi:1.1177/0748175610373459

Enderle, J. L. (1991). *The validation of an assessment instrument for the identification of transition needs: The Enderle-Severson transition rating scale* (Doctoral dissertation). Kentucky: Moorhead State University.

Enderle, J. (2003). *Enderle-Severson Transition Rating Scales*. Moorhead, MN: ESTR Publications.

Falls, J. & Unruh, D. (2010). *Frequently asked questions: Revised part B indicator 14 post-school outcomes.* Retrieved from National Post-School Outcomes Center website: http://www.psocenter.org/content_page_assets/ content_page_19/ind14_faq_rev_may2010.pdf

Finkelstein, V. (2002). The social model of disability repossessed. *Coalition: The Magazine of the Greater Coalition of Disabled People*, 10-16.

Flaugher, R. L. (1978). The many definitions of test bias. *American Psychologist*, *33*, 671.

Fletcher, J. M., Friedman, L., & Pasnak, R. (1998). The validation of the Transition-to-Work Inventory: A job placement system for workers with severe disabilities. *Journal of Developmental and Physical Disabilities*, *10*, 1-22. doi:10.1023/A:1022876232713

French, B. F., & Finch, W. H. (2008). Multigroup confirmatory factor analysis: Locating the invariant referent sets. *Structural Equation Modeling*, *15*, 96-113.

Friedman, L., Cameron, C. C., & Fletcher, J. (1996). *Transition-to-work inventory: A job placement system for workers with severe disabilities*. St. Paul, MN: JIST Works.

Gage, N. A., Lierheimer, K. S., & Goran, L. G. (2012). Characteristics of Students With High-Incidence Disabilities Broadly Defined. *Journal of Disability Policy Studies*, *23*, 168-178.

Gilliam, J. E. (1994). *Work adjustment inventory measures of job-related temperament: Examiner's manual*. Austin, TX: Pro-Ed.

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement*, *37*, 827-838.

Greig, T. C., Nicholls, S. S., Bryson, G. J., & Bell, M. D. (2004). The Vocational Cognitive Rating Scale: A scale for the assessment of cognitive functioning at work for clients with severe mental illness. *Journal of Vocational Rehabilitation*, *21*, 71-81.

Griffin, M. M., Taylor, J. L., Urbano, R. C., & Hodapp, R. M. (2013). Involvement in transition planning meetings among high school students with autism spectrum disorders. *The Journal of Special Education 47,* 1-9.

Guttman, L. (1950). The basis for scalogram analysis. In Stouffer, Guttman, & Suchman (Eds.) *Measurement and prediction*. The American Soldier Vol. IV. New York: Wiley.

Hall, C. W., Peterson, A. D., Webster, R. E., Bolen, L. M., & Brown, M. B. (1999). Perception of nonverbal social cues by regular education, ADHD, and ADHD/LD students. *Psychology in the Schools*, *36*, 505-514.

Hallahan, D. P., & Kauffman, J. M. (1977). Labels, categories, behaviors: ED, LD, and EMR reconsidered. *The Journal of Special Education*, *11*, 139-149.

Harrington, T. F., & O'Shea, A. (2000). *The Harrington-O'Shea career decision-making system revised: Manual*. Circle Pines, MN: American Guidance Services.

Helfrich, C. A., & Fogg, L. F. (2007). Outcomes of a life skills intervention for homeless adults with mental illness. *The Journal of Primary Prevention*, *28*, 313-326.

Hoffman, A., Field, S., & Sawilowsky, S. (2004). Self-determination assessment battery user's guide. *Center for Self-Determination and Transition: Promoting Resiliency and Well-being Throughout the Lifespan.* Detroit, MI: College of Education, Wayne State University. Retrieved from University of Oklahoma website: https://alumni.ou.edu/content/dam/Education/documents/Users%20Guide%20Third%20Edition_1.doc

Holland, J. L., & Rayman, J. R. (1986). The self-directed search. *Advances in vocational psychology*, *1*, 55-82.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1-55.

Hennessey, M. N., Martin, J. E., McConnell, A. E., Terry, R., Martin, J. D., & Willis, D. (2013). *Transition assessment & goal generator technical manual.* Norman, OK: Zarrow Center for Learning Enrichment.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 32, 179-185.

Ihara, E. S., Wolf-Branigin, M., & White, P. (2012). Quality of life and life skill baseline measures of urban adolescents with disabilities. *Social Work In Public Health*, *27*, 658-670. doi:10.1080/19371910903269596

Individuals with Disabilities Education Act (2004). Pub. L. No. 94-142 C.F.R.

International Test Commission. (2001). International guidelines for test use. *International Journal of Testing*, *1*, 93-114.

Ippel, M. J., & Watson, S. E. (2008, September). ASVAB: E pluribus unum. In *50th Annual Conference of the International Military Testing Association,* Amsterdam.

Jastak, J. F., & Jastak, S. R. (1972). *WRIOT-Wide range interest opinion test*. Delaware: Guidance Associates of Delaware, Incorporated.

Johnson, W. L., Mauzey, E., Johnson, A. M., Murphy, S. D., & Zimmerman, K. J. (2001). A higher order analysis of the factor structure of the Myers-Briggs type indicator. *Measurement and Evaluation in Counseling and Development*, *34*, 96-108.

Kauffman, J. M., & Hallahan, D. P. (1995). *The illusion of full inclusion: A comprehensive critique of a current special education bandwagon*. Austin, TX: PRO-ED.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford press.

Knapp, L., & Knapp, R. R. (1982). *Career occupational preference system interest inventory*. San Diego, CA: Edits.

Knapp, L., Knapp, R. R., & Knapp-Lee, L. (1976). *Career ability placement survey*. San Diego, CA: Edits.

Kohler, P. D. (1996). *Taxonomy for transition programming*. Urbana-Champagne: University of Illinois.

Krumbolz, J. D. (1991). *Career Beliefs Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Kuder, F., & Zytowski, D. G. (1991). *Kuder occupational interest survey: Form DD: general manual*. Monterey, CA: CTB Macmillan/McGraw-Hill.

Lachapelle, Y. Y., Wehmeyer, M. L., Haelewyck, M. C., Courbois, Y. Y., Keith, K. D., Schalock, R. R., & ... Walsh, P. N. (2005). The relationship between quality of life and self-determination: An international study. *Journal of Intellectual Disability Research*, *49*, 740-744. doi:10.1111/j.1365-2788.2005.00743.x

Lee, S. H., Wehmeyer, M. L., Palmer, S. B., Soukup, J. H., & Little, T. D. (2008). Self-determination and access to the general education curriculum. *The Journal of Special Education*, *42*, 91-107.

Lee, Y., Wehmeyer, M. L., Palmer, S. B., Williams-Diehm, K., Davies, D. K., & Stock, S. E. (2012). Examining individual and instruction-related predictors of the self-determination of students with disabilities: Multiple regression analyses. *Remedial and Special Education*, *33*, 150-161.

Lewis, T. J. (1994). Full inclusion and the education of children and youth with emotional and behavioral disorders. *Behavioral Disorders*, *19*, 277-93.

Liptak, J. (2012a). *Transition-to-work inventory*. St. Paul, MN: Jist Works.

Liptak, J. (2012b). *Transition-to-work inventory administrator's guide*. St. Paul, MN: Jist Works.

Luckasson, R., Coulter, D., Polloway, E. A., Reiss, S., Schalock, R., Snell, M., ... & Stark, J. (1992). Mental retardation: Definition, diagnosis and systems of supports. *American Psychiatric Association on Mental Retardation*. Washington, DC: American Association of Mental Retardation.

Ludden, L., & Shatkin, L. (2001). *Guide for occupational exploration*. St. Paul, MN: Jist Works.

MacMillan, D. L. (1996). Full inclusion: An empirical perspective. *Behavioral Disorders*, *21*, 145-59.

Marsh, H. W., & Balla, J. (1994). Goodness of fit in confirmatory factor analysis: The effects of sample size and model parsimony. *Quality and Quantity*, *28*, 185-217.

Martin, J. D., Martin, J. E., & Osmani, K. J. (2013). *Teaching Students to Attain Annual Transition Goals Using the Take Action Goal Attainment Lessons*. Career Development and Transition for Exceptional Individuals. doi: 10.1177/2165143413476544

Martin, J. E., Hennessey, M. N., McConnell, A. E., Terry, R., & Willis, D. (2011). *Transition assessment & goal generator.* Norman, OK: Zarrow Center for Learning Enrichment.

Martin, J. E., & Marshall, L. H. (1995). ChoiceMaker: A comprehensive self-determination transition program. *Intervention in School and Clinic*, *30*, 147-56.

McCarney, S. B. (1989). *Transition behavior scale:[technical manual]*. Columbia, MO: Hawthorne Educational Services.

McCarron, L., & Dial, J. G. (1986). *McCarron-Dial evaluation system: A systematic approach to vocational, educational and neuropsychological assessment*. Dallas, TX: McCarron-Dial Systems.

McConnell, A. E., Martin, J. E., Juan, C. Y., Hennessey, M. N., Terry, R. A., el-Kazimi, N. A., ... & Willis, D. M. (2013). Identifying nonacademic behaviors associated with post-school employment and education. *Career Development and Transition for Exceptional Individuals*, *36,* 174-187.

McDougall, J., Evans, J., & Baldwin, P. (2010). The importance of self-determination to perceived quality of life for youth and young adults with chronic conditions and disabilities. *Remedial And Special Education*, *31*, 252-260.

McGuire, J., & McDonnell, J. (2008). Relations between recreation and levels of self-determination for adolescents and young adults with disabilities. *Career Development for Exceptional Individuals*, *31*, 154-163.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal Of Applied Psychology*, *93*(3), 568-592.

Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, *7*, 361-388.

Mesibov, G. B., & Shea, V. (1996). Full inclusion and students with autism. *Journal of Autism and Developmental Disorders*, *26*, 337-346.

Messick, S. J. (1989). Validity. In Linn, R.L. (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*, 111-13.

Mithaug, D. E., Campeau, P. L., & Wolman, J. M. (2003). Assessing self-determination prospects among students with and without disabilities. In *Self-determined learning theory: Construction, verification, and evaluation*, (pp. 61-76). Mahwah, NJ: Lawrence Erlbaum Associates.

Mosier, C. I. (1947). A critical examination of the concepts of face validity. *Educational and Psychological Measurement, 7,* 191-205.

Myers, J. B. & Briggs, K. C. (1988). *Myers-Briggs Type Indicator Form M.* Palo Alto, CA: Consulting Psychologists Press.

National Secondary Transition Technical Assistance Center (2011). *Age appropriate transition assessment toolkit third edition*. Charlotte, NC: University of North Carolina.

Neubert, D. A., & Leconte, P. J. (2013). Age-appropriate transition assessment: the position of the division on career development and transition. *Career Development and Transition for Exceptional Individuals,* 36, 72-83.

Nollan, K. A., Horn, M., Downs, A. C., Pecora, P. J., & Bressani, R. V. (2002). *Ansell-Casey life skills assessment (ACLSA) and life skills guidebook manual*. Seattle, WA: Casey Family Programs.

O'Connor, C., & Fernandez, S. D. (2006). Race, class, and disproportionality: Reevaluating the relationship between poverty and special education placement. *Educational Researcher*, *35*, 6-11.

Oliver, M. (2004). If I had a hammer: The social model in action. *Disabling barriers–Enabling environments*, *2*, 7-12.

Osipow, S. H., Carney, C. C., Winer, J., Yanico, B., & Koschier, M. (1987). *Career Decision Scale*. Odessa, FL: Psychological Assessment Resources.

Palmer, S. B., & Wehmeyer, M. L. (2003). Promoting self-determination in early elementary school: Teaching self-regulated problem-solving and goal-setting skills. *Remedial and Special Education*, *24*, 115-126.

Parker, R. M. (1983). *Occupational aptitude survey and interest schedule: Aptitude survey*. Austin, TX: Pro-Ed.

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P., & Fleishman, E. A. (1999). *An occupational information system for the 21st century: The development of O* NET*. Washington, DC: American Psychological Association.

Pirritano, M. J. (2005) Gamma Hat and McDonald's NCI Calculator [MS Excel Worksheet Formulas].

Popham, W. (2008, September). A misunderstood grail. *Educational Leadership*, 82-83.

Prien, E. P., & Hughes, G. L. (1993). *Job analysis and ADA compliance*. A workshop presented at the 1993 Society for Industrial and Organizational Psychology, San Francisco, CA.

Primé, D. R., & Tracey, T. J. (2010). Psychometric properties of the career clusters interest survey. *Journal of Career Assessment*, *18*, 177-188.

Rehabilitation Act of 1973, 29 U.S.C. § 701 et seq.

Riordan, C. M., & Vandenberg, R. J. (1994). A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management*, *20*, 643-671.

Rosinek, M. (1984). *Prevocational assessment screen (PAS) evaluator's manual*. Washington, DC: Distributed by ERIC Clearinghouse.

Saal, F. E., Downey, R. G., & Lahey, M. A. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin*, *88*, 413-428.

Salvia, J., Ysseldyke, J. E., & Bolt, S. (2004). *Assessment in special and inclusive education*. Boston, MA: Houghton Mifflin.

Sampson, J. P., Peterson, G. W., Lenz, J. G., Reardon, R. C., & Saunders, D. E. (1996). *Career thoughts inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350.

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, *18*, 210-222. doi:10.1016/j.hrmr.2008.03.003

Semrud-Clikeman, M., Walkowiak, J., Wilkinson, A., & Christopher, G. (2010). Neuropsychological differences among children with Asperger syndrome, nonverbal learning disabilities, attention deficit disorder, and controls. *Developmental neuropsychology*, *35*, 582-600.

Shogren, K. A., Wehmeyer, M. L., Palmer, S. B., Soukup, J., Little, T. D., Garner, N., & Lawrence, M. (2008). Understanding the construct of self-determination: Examining the relationship between the ARC's self-determination scale and the American institute for research’s self-determination scale. *Assessment for Effective Intervention*, *33*, 94-107.

Shrive, F. M., Stuart, H., Quan, H., & Ghali, W. A. (2006). Dealing with missing data in a multi-question depression scale: A comparison of imputation methods. *BMC Medical Research Methodology*, *6*, 57-67.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107-120.

Smith, T. L., Polloway, E. A., Smith, J., & Patton, J. R. (2007). Self-determination for persons with developmental disabilities: Ethical considerations for teachers. *Education And Training In Developmental Disabilities*, *42*, 144-151.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scales: Interview edition, survey form manual*. Circle Pines, MN: American Guidance Service.

Stein, M. T., Klin, A., & Miller, K. (2004). When Asperger's syndrome and a nonverbal learning disability look alike. *Pediatrics*, *114*(Supplement 6), 1458-1463.

Streiner, D. L. (2003). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of Personality Assessment*, *80*, 217-222.

Sternberg, R. J., & Hedlund, J. (2002). Practical intelligence, g, and work psychology. *Human Performance*, *15*(1-2), 143-160.

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational measurement*, *27*, 361-370.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson.

Taylor, B. N. (2009). *Guidelines for evaluating and expressing the uncertainty of NIST measurement results*. Darby, PA: DIANE Publishing.

Test, D. W., Mazzotti, V. L., Mustian, A. L., Fowler, C. H., Kortering, L., & Kohler, P. (2009). Evidence-based secondary transition predictors for improving postschool outcomes for students with disabilities. *Career Development for Exceptional Individuals*, *32*, 160-181.

Thoma, C. A., Rogan, P., & Baker, S. R. (2001). Student involvement in transition planning: Unheard voices. *Education and Training in Mental Retardation and Developmental Disabilities*, *36*, 16-29.

Trainor, A. A., Patton, J. R., & Clark, G. M. (2005). *Case Studies in Assessment for Transition Planning*. Austin, TX: Pro-Ed.

Valsiner, J., & Van der Veer, R. (1999). The encoding of distance: The concept of the zone of proximal development and its interpretations. *Lev Vygotsky: Critical assessments*, *3*, 3-31.

Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, *5*, 139-158.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, *3*, 4-70.

Vandercook, T. (1989). The *McGill Action Planning System* (*MAPS*): A strategy for building the vision. *Journal of the Association for Persons with Severe Handicaps (JASH)*, *14*, 205-15.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *31*, 321-327.

Waldron, N. L., & McLeskey, J. (2010). Inclusive school placements and surplus/deficit in performance for students with intellectual disabilities: Is there a connection? *Life Span and Disability*, 29.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for children, revised*. San Antonio, TX: Psychological Corporation.

Wehmeyer, M. L. (1995). *The Arc's self-determination scale: Procedural guidelines*. Washington, D.C.: *ARC* of the US.

Wehmeyer, M. L., Agran, M., & Hughes, C. (1998). *Teaching self-determination to students with disabilities: Basic skills for successful transition*. Baltimore, MD: Paul H. Brookes Publishing.

Wehmeyer, M. L., Field, S., Doren, B., Jones, B., & Mason, C. (2004). Self-determination and student involvement in standards-based reform. *Exceptional Children*, *70*, 413-425.

Wehmeyer, M. L. & Kelchner, K. (1995). *The* ARC*'s self-determination scale: Procedural guidelines*. Washington, D.C.: *ARC* of the US. Retrieved from University of Oklahoma Zarrow Center website: http://www.ou.edu/content/education/centers-and-partnerships/zarrow/self-determination-assessment-tools/arc-self-determination-scale.html

Wehmeyer, M., & Schwartz, M. (1997). Self-determination and positive adult outcomes: A follow-up study of youth with mental retardation or learning disabilities. *Exceptional Children*, *63*, 245-256.

Wiesen, J. P. (1997). *Technical manual for the wiesen test of mechanical aptitude (WTMA)*. Newton, Massachusetts: Applied Personnel Research.

Williams, D. L., Goldstein, G., Kojkowski, N., & Minshew, N. J. (2008). Do individuals with high functioning autism have the IQ profile associated with nonverbal learning disability?. *Research in autism spectrum disorders*, *2*, 353-361.

Wolman, J., Campeau, P., Dubois, P., Mithaug, D., & Stolarski, V. (1994). *AIR self-determination scale and user guide*. Palo Alto, CA: American Institute for Research.

Ysseldyke, J. E., Algozzine, B., Richey, L., & Graden, J. (1982). Declaring students eligible for learning disability services: Why bother with the data? *Learning Disability Quarterly*, *5*, 37-44.

Ysseldyke, J. E., Algozzine, B., Rostollan, D., & Shinn, M. (1981). A content analysis of the data presented at special education placement team meetings. *Journal of clinical psychology*, *37,* 655-662.

Zhang, D. (2001). Self-determination and inclusion: Are students with mild mental retardation more self-determined in regular classrooms? *Education & Training In Mental Retardation & Developmental Disabilities*, *36*, 357-362.

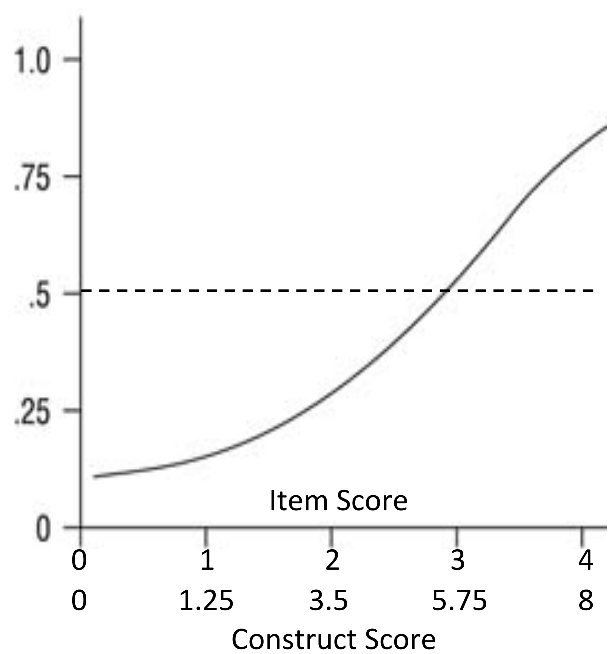
Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF).* Ottawa: National Defense Headquarters.

Appendix A

The *TAGG* uses item response theory (IRT) to generate each student's score. IRT uses estimation to assess the probability that a test will obtain a particular score. The *TAGG* developers considered an item "mastered" if the test taker had a 50% probability of scoring a three or higher on the professional or family version Likert-type items, or a one or higher on student version Likert-type items.

The figure below provides a visual representation the probability that a student having a particular construct score will be rated at a certain level on an item. The x-axis represents the range of item scores for the professional version of the *TAGG* (0-4) *and* construct scores. The y-axis represents the probably that a student at a construct level will be given the score. The dashed horizontal represents the 50% probably that a student will obtain a certain score.

Probability that a student will have mastered *TAGG* Professional version item 10: Student keeps working until he has accomplished a goal.



Based on this graph, a student who has about a 5.75 score on the Persistence construct has a 50% probability of having a score of three on item 10. Therefore, any test taker who scores a 5.75 or greater in persistence is believed to have mastered item 10, while students who have a score less than 5.75 are not believed to have mastered the item.