The Forgotten: Formal Assessment of the Adult Writer

Daniel J. McNair
Toi L. Curry
Georgia Southern University

Abstract

This review of current writing assessment practices focuses upon the adult population, an area significantly underrepresented within psychoeducational literature. As compared to other populations, such as K-12 students, there are few options for the practitioner wishing to evaluate adult writers by means of standardized assessment instruments. This review of literature discusses common approaches to written expression assessment. While indirect assessment and cognitive processing approaches are reviewed in brief, significant attention is given towards the traditional approach of direct writing assessment. Aspects of direct assessment methods include syntax, cohesion, sense of audience, spelling, and fluency. In addition to these factors, the role of story prompts, rater reliability, and affective variables are considered. The paper concludes with brief overviews of applicable standardized assessment instruments for written expression assessment of this population. Particular focus is given towards the Wechsler Individual Achievement Test-II (WIAT-II) and the more recent edition of the Wechsler Individual Achievement Test-III (WIAT-III). The authors contend that although the WIAT-II may have been an adequate instrument for direct assessment of writing ability in adults, test developers have failed to maintain a focus on this population in the more recent WIAT-III. This inadequacy is evidenced in the omission of grade-based scoring tables for college populations and lack of content appropriate writing prompts for adults. Implications for practitioners working with postsecondary populations are discussed.

Keywords: Postsecondary, assessment, written expression, learning disability

Written expression simply refers to the ability to communicate one’s thoughts and feelings through the written word; however, this activity is no simple process. Unlike basic writing skills such as handwriting, spelling, or sentence composition, written expression is a more involved process that requires the use of both basic writing and more complex cognitive skills. Although written composition requires mastery of elementary processes noted above, cognitive skills such as planning, organization, and cohesion are also required to create meaningful text passages for effectively communicating with others (Katz, Goldstein, & Beers, 2001). Not only is writing ability a valuable life skill, but the development of such ability is important for determining an individual’s ability to navigate educational systems, work environments, and social situations (Cole, Muenz, Ouchi, Kaufman, & Kaufman, 1997). As noted by Gregg, Coleman, and Lindstrom (2008), as the societal demands of literacy increase, writing ability becomes an increasingly important factor in allowing an individual to graduate from high school, achieve in the postsecondary setting, and succeed in the work environment. Many State exams require successful completion of writing assessments. Gregg et al. (2008) also note that the recent inclusion of an essay section on the College Board’s SAT I test provides further indication of the increased concern of writing ability in graduating high school students.

Students transitioning from the secondary to postsecondary educational setting encounter multiple obstacles that include, but are not limited to, time management difficulties, increased stress, acclimation to a different environment, and new instructional methods. The list of potential difficulties for college students is non-exhaustive. However, in addition to these factors, students with disabilities also encounter unique obstacles that may occur as a result of their transition from the service model of K-12 educational systems.
According to Gregg (2009), after some frustration and much confusion, many students with disabilities learn that the legal requirements of disability documentation change in this new educational setting. Documentation requirements for learning disabilities in the postsecondary setting are often more comprehensive, therefore many students learn that their current eligibility documentation does not meet legal requirements for receipt of services. Therefore, there are students who must engage in additional psychoeducational assessments at the postsecondary level to provide adequate documentation of learning difficulties that were previously accommodated before entering tertiary educational settings. These psychoeducational evaluations are generally more stringent and encompassing measures of functioning, for which writing ability is a necessary component (Gregg, 2009). However, the assessment of writing ability using normative measures proves difficult for the practitioner working with adult populations. In comparison to other areas commonly assessed by psychologists, there is a lack of available instruments for the assessment of written expression in this population.

Evaluating the Adult Writer

The evaluation of writing ability is not a new concept. While models of written language disorders have been discussed by researchers beforehand, it was not until the 1970’s that the diagnosis and treatment of written language disorders were legislated by the U.S. Office of Education (Hooper, 2002). However, despite this history of research, many professionals continue to lack confidence or understanding of the diagnosis, treatment, and definition of written language disabilities (Gregg, 1995). According to Hooper (2002), empirically-based research of written language disorders has only begun to gain increased popularity in the past two decades. In addition, these studies have not contributed greatly towards the understanding of etiology, developmental progression, or effective treatment for written expression disorders (Katz et al., 2001).

The lack of available research, in part, may be due to the nature of the writing process as a whole. To define a written language disability in and of itself poses difficulty since written language is such a broad term that encompasses so many skills and domains of ability (Cole et al., 1997; Katz et al., 2001). For example, the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision ([APA] American Psychiatric Association, 2000) defines a disorder of written expression as writing skills that are substantially below expected levels, given chronological age, intelligence, and education. An additional criteria is that these writing skill disturbances must also interfere with either academic progress or daily living. The next version of the DSM is currently in development with an expected release date of spring 2013. Preliminary reports suggest that learning disabilities will be grouped into one classification with specifiers rather than separate categories for each academic area. At the time of this writing, it also appears that diagnostic criteria will be reflective of those set forth by the Reauthorization of the Individuals with Disabilities Education Improvement Act of 2004 (IDEA). In other words, the DSM-V is expected to provide a more structured approach to diagnosing learning disabilities than currently exposed by the 4th edition of this manual. Although these definitions do provide some qualifying data for diagnosis, authors acknowledge that when compared to other learning disorders, there is less understanding of written expression disability and the standardized assessments of writing are not as well developed as those measuring reading or mathematical abilities. For the purposes of this paper, the terms “disorder of written expression” and “written expression disability” are used interchangeably.

The lack of research also creates uncertainty for practitioners wishing to evaluate the writing ability of individuals. Although the quantity of research on writing disorders has grown significantly over recent decades, the primary focus of previous researchers appears to have been on the writing of children (Newcomer & Barenbaum, 1991). Specifically, Newcomer and Barenbaum provide a comprehensive overview of studies on the writing abilities of children with learning disabilities occurring between 1980 and 1990. During this decade, studies of written expression appear to have not only increased in quantity, but also undergone a transitional shift from focusing upon basic factors such as syntax, fluency, and mechanics toward the more complex analysis of text structures, metacognition, and response to practice or intervention (Newcomer & Barenbaum, 1991).

With the exception of a few researchers, the majority of current learning disability studies continue to focus upon younger populations and fails to adequately address assessment and intervention for the writing of adult age populations (Gregg, Coleman, Davis, Lindstrom, & Hartwig, 2006; Gregg, Coleman, Stennet, &
It is likely that this underrepresentation may be attributed towards the increased focus upon early identification and treatment for learning disorders in school-aged children. Naturally, more attention has been given towards offering assistance to these younger populations with the enactment of legislation such as IDEA (2004) and the No Child Left Behind Act of 2002 (NCLB). In fact, most funding initiatives and programs have been directed toward the early identification and treatment for younger students with the supposition that earlier intervention will thwart or eradicate difficulties in later years (Gregg, 2009). However, researchers have illustrated that although such services are effective, individuals with learning disabilities often encounter continued difficulties despite such early attempts at remediation (Gregg, 2009).

While NCLB (2002) and the Reauthorization of IDEA (2004) are intended for K-12 settings, college age populations are served by means of Section 504 of the Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990. Section 504 was designed to protect the civil rights of individuals with disabilities in programs and activities receiving federal funding. Widely recognized as the first civil rights statute for individuals with disabilities, Section 504 ensures that persons are not excluded from participation or denied benefits of public services based solely upon reason of a disability (Lissner, 1997). The Americans with Disabilities Act ([ADA], 1990), while similar to Section 504, expands upon the previous federal law by broadening the agencies and businesses that must adhere to nondiscriminatory operations. The ADA was amended in 2008 to provide further clarification of those protections described in the original law. Much of the language in each of these federal laws is similar in that they guarantee civil rights protections to individuals with disabilities. These civil rights are similar to those provided to individuals that protect against discrimination on the basis of race, color, sex, nationality, age, or religion (Lissner, 1997; Mangrum & Strichart, 1988).

There are, however, great differences in service delivery models resulting from these important legislative rulings. Whereas practitioners assessing school-aged children spend a great deal of time recommending interventions and guiding intervention delivery in order to increase skills and proficiencies, evaluations of college-age populations are primarily conducted to provide the documentation that is needed for an individual to gain access to accommodations (Gregg et al., 2006). In effect, whereas school age children are “treated” for a disability through services and interventions, the postsecondary student is granted “equal access” to the curriculum through accommodations, but there is no legislative requirement for the individual to receive remediation (Gregg et al., 2006). In other words, the primary purpose of IDEA (2004) and other K-12 legislation has been to provide remedial services to students and the ADA (1990) primarily serves the purpose of granting equal opportunity to students with disabilities in the postsecondary setting. Since relatively little research exists pertaining to college-aged learners who are writing disabled, study findings from younger populations are often generalized towards the assessment of college age populations. Although these studies with younger populations do provide valuable information pertaining to the identification of basic writing processes, they do not replace the need for research into the writing processes of older individuals.

**Learning Disabilities and Typical Development**

A common topic in writing disorder research is in differentiating between writers with learning disabilities and those writers who are typically developing. Numerous studies outlining characteristics of students with writing disabilities include analysis of word counts of written passages, diversity of word choices, and the inclusion of words with more than two syllables (Gregg et al., 2002).

In a study comparing the textual cohesion of college-aged students with and without diagnosed learning disabilities, researchers analyzed the expository writing of 45 female college students using an adapted scale for measuring the cohesive constructs of grammar ties, transition ties, and lexical ties (Gregg, 1985). Researchers utilized this adapted scale to obtain frequency and accuracy counts of these cohesive structures and concluded that students experiencing significant writing deficits encounter difficulties in their use of structural cohesion when writing. While overall results did not demonstrate significant differences between cohesive ties, findings did suggest that the students with learning disabilities were more likely to commit errors in accuracy, morphological endings, and omission of words (Gregg, 1985). Additionally, Gregg (1985) suggested that students with writing disabilities use fewer demonstratives in written language and may exhibit a lack of flexibility and diversity in word choices. This
reluctance to utilize more complex transitional ties is often more reflective of inexperienced writers (Gregg, 1985; Gregg et al., 2008).

In a review of an elementary and middle school writing research program, investigators’ efforts were targeted towards differentiating between students with learning disabilities and students without learning disabilities (Graham, Harris, MacArthur, & Schwartz, 1991). After reviewing their findings from four previous studies, the group concluded that typically developing writers were more likely to create responses that convey meaningful thoughts and ideas and were also more likely to focus on redeveloping their ideas during the revision process. The students with learning disabilities were found to be more likely to focus upon correcting basic grammar or mechanical errors rather than the substance of their written product. Interestingly, each group of writers appeared to demonstrate the same levels of confidence in their own writing abilities (Graham et al., 1991).

In a recent study focusing on the perceived writing quality of students with a writing disability and students without a writing disability, researchers suggest that the proper use of organization and cohesion are just as important as punctuation, grammar, and spelling (Gregg, Coleman, Davis, & Chalk, 2007). To reach these conclusions, researchers provided raters with three products of each student’s writing sample. The first product was the original handwritten sample. The second was a typed version. The third version was typed and edited for spelling, punctuation, and grammatical errors without altering word complexity or cohesive ties. Overall results of this study concluded that the qualitative perceptions of raters were significantly influenced by the quantity of spelling errors contained in a writing sample. Handwriting ability and illegible writing samples also appeared to play a significant role in the scoring process. Both handwriting ability and spelling negatively influenced rater perceptions of a student’s samples. In addition, the participants with dyslexia also received lower scores than the typically developing group when evaluators assessed the third version of writing that had been typed and corrected for spelling, punctuation, and grammar. Overall results suggest that even though handwriting and spelling play an important role in the student’s ability to receive quality writing scores, the ability to succinctly and accurately convey a message in an organized and cohesive sample of writing may be just as important (Gregg et al., 2007).

Approaches to Assessment

Direct and Indirect Assessment Methods

There are two major theoretical methods for assessing the writing ability of individuals. These approaches are often utilized together in qualitative assessment, but formal standardized written expression assessments often focus upon one or the other. These approaches are termed “direct” and “indirect” assessment. Direct assessment refers to those methods of evaluation in which the learner is required to generate complete written discourse or essay (Hooper, 2002). These types of evaluations require a prompt such as a story starter or visual cue, and the writer then creates a written composition to meet task demands set forth by the prompt.

The other approach, termed indirect assessment, requires the learner to accurately respond to questions that target basic writing skills such as grammar, mechanics, and spelling (Hooper, 2002). The indirect approach does not require the learner to create a sample of connected discourse, but the goal of this approach is to identify skill areas and knowledge of writing conventions and rules without the requirement to apply those skills in creating an original sample of writing. The indirect approach most often consists of multiple choice formats, single item responses, and brief sentence composition (Muenz, Ouchi, Cole, 1999). The results of indirect assessment methods are then generalized to the broader skill of written expression ability.

Researchers have long debated the validity and reliability of each of these methods (Breland & Gaynor, 1979; Hooper, 2002; Miller & Crocker, 1990; Sabban & Kay, 1987). Miller and Crocker (1990) attest to this level of controversy and claim that the debate surrounding writing assessment has been greater than that of any other area of academic functioning. In addition, these researchers explain that although reading and math assessments have played a critical role in the development of many assessment programs, the development of writing assessments have not benefitted from the same degree of focus (Miller & Crocker, 1990).

Each of these approaches to writing assessment has advantages and disadvantages. Generally speaking, proponents of direct assessment argue that such methods provide a face valid artifact of writing ability (Breland & Gaynor, 1979; Miller & Crocker, 1990; Powers, Fowles, & Willard, 1994; Sabban & Kay, 1987). Supporters of indirect assessment methods
argue that such approaches are more reliable and therefore a more accurate measure of writing skill. Even though there is no writing product, indirect assessment results are more objective and thus provide greater reliabilities than those produced in the qualitative assessment of writing samples (Gregg, 1995; Gregg et al., 2007; Muenz et al., 1999). Miller and Crocker (1990) further their argument for indirect methods by citing the easier-to-score format of multiple choice or short answer responses over the often complicated and labor intensive process of analyzing a body of text.

Cognitive Processes of Writing

In addition to the assessment of writing through direct or indirect methods, researchers and practitioners have also identified multiple cognitive processes that are associated with writing skill. Executive functioning, crystallized intelligence, working memory, and long-term storage and retrieval have each been identified as basic cognitive processes that are necessary for proficient and skilled writing ability (Alamargot, Caporossi, Chesnet, & Ros, 2011; Gregg, 1995; Gregg et al., 2007; Hillis, 2008; MacArthur, Graham, & Fitzgerald, 2008). Identification of weaknesses in these cognitive areas may provide the practitioner with additional information when evaluating adult learners. But, as expected, research into the neuropsychological processes involved in the writing of college populations is lacking (Newcomer & Barenbaum, 1991; Semrud-Clikeman & Harder, 2010).

However, a recent study does provide some interesting findings concerning adult writers and working memory capacity (Alamargot et al., 2011). Although previous studies involving children have demonstrated the role of working memory in more elementary processes such as lexicon or orthographic processing, this study suggests that, for adults, working memory capacity is instrumental to the higher demands of the writing process such as audience awareness. These researchers evaluated the influence of working memory capacity on communicative efficiency and sense of audience in a procedural writing task. To achieve this, the eye and graphomotor movements of 25 graduate students were tracked using computer software during a writing task. The students were asked to create a written instructional guide for assembling a model turbine. Using the computer software, researchers were able to track the frequency of pauses and references toward a visual diagram provided to each writer to assist in their composition. By requiring the students to write about a complex procedural process, each writer had to refer back to the diagram. Findings suggest that the ability to maintain a sense of audience or a visual representation in memory while writing is a direct reflection of an individual’s working memory capacity. Researchers also postulate that those individuals with stronger working memory were able to better communicate with their intended audience during the task (Alamargot et al., 2011).

Nevertheless, the majority of available literature on cognitive processes and writing ability does pertain to children. Therefore, practitioners must reference the more available literature discussions of children’s writing abilities when assessing the writing of adults. However, researchers forewarn that all research into the writing ability of children should not be overgeneralized toward adult writers since many variables such as cognitive processes, language development, and experiences may vary across the lifespan (Gregg et al., 2008). Therefore, the evaluation of an adult’s writing ability should not be solely based upon the cognitive processing literature as it relates to written expression in children. For this reason, either direct or indirect assessment approaches remain the most fruitful methods for determining the presence of a writing disorder in adult populations.

Direct Assessment Methods

Multiple areas have been identified as important factors to consider when evaluating written expression ability. Current research suggests the evaluation of syntax, cohesion, sense of audience, spelling, and fluency are instrumental in the determination of a writing deficit (Gregg et al., 2008).

Syntax is defined as the number of words, clauses, diversity of sentence structures, variety of word choices, and error frequencies within a writing sample (Gregg et al., 2002; Gregg et al., 2008). These figures provide the evaluator with measures of quantity, accuracy, and diversity in a writing sample. In fact, it has been noted that a majority of research into the writing of college-aged populations has focused upon “frequency counts” of syntax components such as error count or number of sentences (Gregg et al., 2002).

Cohesion refers to the correct usage of cohesive ties and transitions (Gregg, 1995; Gregg et al., 2008). A passage lacking appropriate cohesion may rely upon restricted word choices or may contain ambiguous
pronouns or demonstratives. Other significant effects of poor cohesion include a lack of organization or a passage that does not flow (Gregg, 1995; Gregg et al., 2008).

Sense of audience not only refers to the ability of the writer to identify the audience of a composition but also the ability to evaluate the purpose of a writing activity correctly (Gregg et al., 2008). Failure to consider possible readers may be a common occurrence for underachieving writers and such negligence on the part of the writer may result in ambiguous references or omission of supporting details. Also of importance and closely related to sense of audience is the ability of a writer to evaluate and perform edits to a writing sample to better address the needs of an intended audience. For the individual, the ability to take on the reader’s perspective is an important skill when reviewing, editing, and revising a sample of writing (Gregg et al., 2008).

Spelling difficulties may also greatly affect a writer’s ability to communicate ideas effectively. While spelling errors may occasionally influence the meaning of discourse, the lack of automaticity that is associated with poor spelling may play a more critical role in the quality of writing (Gregg & Mather, 2002). Since poor spellers may need to pause due to a lack of automaticity in spelling words, they may lose track or fail to remember planned ideas or concepts (Gregg & Mather, 2002; Viel-Ruma, Houchins, & Fredrick, 2007). In addition, the compositions of poor spellers may lack the diversity of word choices that is present in the writing of better spellers. This failure to utilize a greater complexity of word choices may portray the poor speller’s writing as less sophisticated or scholarly (Viel-Ruma et al., 2007). Previous research has also demonstrated the positive correlation between spelling ability and writing quality (Viel-Ruma et al., 2007).

The quantity of written text is another important consideration when evaluating a sample of writing. Although substance is important, a writer’s ability to work efficiently also plays a critical role in the writing process. Fluency refers to the quantity of words or sentences a writer is able to produce under timed conditions. In other words, fluency is simply a measure of the amount of writing that may be produced in a given time limit (Gregg et al., 2007; Gregg, et al., 2008).

Previous research has also suggested that the particular modality - such as pencil and paper, computer, or voice activated software - may also play a role in the quality and quantity of writing for some individuals. For example, the use of a computer during the writing process has been shown to improve fluency, editing, and quality of writing for both students with learning disabilities and typically developing students (Gregg et al., 2008; Gregg, 2009). According to Gregg (2009), the use of computers to accommodate writing difficulties has increased significantly since the 1980’s. Other tools to accommodate writing difficulties include additional time, scribe/voice recorder, voice to text software, and word prediction software (Gregg, 2009).

**Story Prompts**

The content of a writing sample may be greatly influenced by the level of knowledge a writer possesses about a given topic; therefore, a writer may be able to produce more elaborate and better-organized text when working with a familiar topic (Gregg et al., 2008). Not only does a writer’s product appear to be influenced by topic knowledge but also by the type of prompt utilized in a writing evaluation. Prior research reveals that writer productivity may be greatly influenced by story prompts and directives.

Previous research by Hooper et al. (1994) contends that a story prompt portraying at least two characters, an interesting scene, and some type of potential conflict by means of a color photograph may elicit the greatest possible response by writers. Other researchers have tested the hypothesis of Hooper et al. (1994) by evaluating the text of writers after exposure to two very different pictorial stimuli (Cole et al., 1997). A sample of 50 individuals, aged 13-46 years, was administered a stick figure pictorial stimulus and a photographic stimulus in two separate writing activities to ascertain the degree to which pictorial prompts affect quality of writing. The stick figure stimulus was taken from the PIAT-R written expression level II subtest and the photograph stimulus was created to fulfill all four criteria set forth by Hooper et al. (1994). An individualized scoring system was adapted from the PIAT-R and WIAT written expression scoring systems to properly evaluate for structural and mechanical characteristics of samples. Trained raters independently scored each sample; resulting scores between pictorial stimuli were significantly different. While mechanical items such as grammar and punctuation were unaffected, the compositions differed markedly in their use of structural items such as unity, organization, and development of ideas. Researchers concluded that the type of prompt used in
a writing assessment makes a critical difference in the resulting writing sample (Cole et al., 1997).

Expanding upon prior work, Muenz, Cole, and Balderson (2000) evaluated the potential effects of a rater’s prior knowledge of research hypotheses in evaluations of “Hooper-like and non-Hooper-like” pictorial stimuli. In a study mirroring the work by Cole et al. (1997), a smaller sample of 29 participants aged 11-14 years were evaluated. All research parameters were replicated with the one exception of the rater who, in this instance, was unaware of the research hypothesis. As previously ascertained in the larger study, there was a significant difference observed in scores of the photographic stimuli meeting Hooper et al. (1994) criteria as compared to text resulting from the stick figure stimulus (Muenz et al., 2000). These findings provide further support for the criteria set forth by Hooper et al. (1994) and account for any possible rater bias that may have occurred in the initial study of larger size.

**Affective Variables**

Other considerations in the assessment of a writing sample include situational or affective variables of the writer. Writing performance may be influenced by the socio-emotional state of a writer and writing anxiety is a well-documented phenomenon that has garnered significant attention by educators and researchers. Previous research has shown that students with increased anxiety may receive lower grades on essays, written assessments, and writing tests (Martínez, Koch, & Cass, 2011).

Writing anxiety may be attributed to dispositional attitudes accumulated over time and prior experiences or it may be a result of a specific activating event such as the requirement of a specific writing task (Martínez et al., 2011). According to Ucgun (2011), writing anxiety may begin during early childhood and continue throughout the lifespan. Manifestations or presentations of writing anxiety may include tension, procrastination, physiological symptoms, or preoccupation (Martínez et al., 2011). Generally, the written products of anxious writers tend to be shorter in length, less fluent, and less diverse in selection of word choices (Ucgun, 2011).

Researchers have postulated that writing anxiety or apprehension may be related to the amount of writing experiences of an individual. In a study of college-aged students, 127 individuals participated in pre- and post-semester surveys to gauge writing anxiety and self-efficacy. Findings of the study conclude that as writing experiences increase, associated anxiety of writing decreases (Martínez et al., 2011). In other words, the more an individual is exposed to writing activities, the greater the likelihood that person may overcome writing anxieties. Results of the study also indicate that students demonstrating avoidance behaviors are more likely to experience additional writing anxiety over time (Martínez et al., 2011).

Further evidence of the benefits of increased writing exposure is provided by Shweiker-Marra and Marra (2000) in a study of 29 at risk 5th grade students. These students, identified as struggling writers, participated in a writing program to determine the effect of increased exposure to prewriting activities as a means of decreasing writing anxiety. Results of the study suggest that participants were able to decrease their writing anxiety and improve overall writing performance through increased exposure to prewriting activities (Schweiker-Marra & Marra, 2000).

In a study of the writing anxiety of elementary age students, researchers surveyed 1,407 Turkish students and found that writing anxiety decreased as a result of increased exposure to writing activities. Additionally, study results suggest that students exhibiting lower levels of writing anxiety were more likely to enjoy language classes, keep personal diaries, and read more books (Ucgun, 2011).

Available research suggests that affective variables such as anxiety may influence a writer’s ability to construct written discourse that represents his or her best effort. In addition, it appears that a lack of positive writing experiences may be a contributing factor to the development of writing anxiety. Gregg and Mather (2002) have accepted the ramifications of affective variables to writing to such a degree that they encourage the use of a writing apprehension or writing anxiety scale when a writer’s product appears to be influenced by such factors.

**Reliability of Ratings**

As previously mentioned, a significant concern of direct writing assessments is that of reliability. Even though the required writing activity and normative sample can control a degree of variability in scores, the greatest threat to test reliability is that of ratings given to samples of writing. Most test developers account for this reliability issue by providing some form of a scoring rubric.
Noted in a study conducted by Muenz et al. (1999), most forms of direct writing assessment have either poor reliability (coefficients less than .70) or lack appropriate validity. In this specific study, researchers compared the reliability and validity of items contained within the Peabody Individual Achievement Test-Revised (PIAT-R) and Wechsler Individual Achievement Test (WIAT). In their analysis, these instruments were administered to 50 individuals aged 13 to 46 years. A panel of three graduate students conducted ratings, while reliability of items were established by means of inter-rater agreement and item validity was determined by item-total correlations. Overall results of the study indicated that the WIAT Written Expression Test (WIAT) provides a greater number of reliable and valid items than the PIAT-R (Muenz et al., 1999). However, additional research findings from this study may have more importance for the contemporary assessment of written expression. During the course of their study, researchers made an important distinction between structure and mechanics. Structural items - those that assess quality through means of unity, cohesion, organization, and idea development - were found to be more reliable amongst raters than the objective measurement of mechanics. Mechanics include the direct measures of grammar, punctuation, and spelling. Although authors expected that mechanics (grammar, punctuation, etc.) would be more reliable due to the objective rules for language use, the structural items were found to be more reliable between raters. Authors postulated that this unexpected result was due to the tendency for significant errors in writing structure such as cohesion or organization to be more readily apparent to readers on a consistent basis. Mechanical errors such as grammar and punctuation mistakes may not have been consistently identified if raters were relying upon memory for such rules when scoring writing samples. As a result, researchers concluded that future measures of written expression should include the use of a comprehensive reference source in addition to a scoring system of greater scope that would allow for greater distinctions between subtle differences in writing.

In an analysis of two distinct written expression rating systems, Knoch (2009) gathered score results from 10 evaluators after they rated 100 samples of writing. The raters utilized an empirically developed scale containing specific descriptors and an intuitively developed scale containing less specific descriptors to score each sample. After analysis of results, Knoch (2009) concluded rater reliability to be substantially greater when more comprehensive guidelines and descriptors are illustrated within scoring guidelines.

The benefit of training raters to interpret scoring guidelines was investigated in a study of primary grade instructors (Stuhlmann, Daniel, Dellinger, Kenton, & Powers, 1999). Researchers first split a group of 40 instructors into groupings of 23 and 17 individuals. The larger group was then trained to interpret a scoring rubric, while the smaller group of individuals received no training for use of the rubric scoring system. Each group then independently scored 20 first-grade writing samples based upon guidelines set forth by the scoring rubric. Resulting data indicated there to be more variability in assigned scores for the untrained rater group as compared to that of the trained group. Researchers conclude that training raters in their use of predetermined scoring guidelines increases their ability to reliably rate writing samples (Stuhlmann et al., 1999).

Assessment Instruments

There are few standardized writing assessments that are applicable to the college or adult population. While there are many choices for the evaluation of persons under age 18, there are limited options for the assessment of an adult writer. There are a few instruments that may be used for young adults, such as the Kaufman Test of Educational Achievement-II (KTEA-II), the Oral and Written Language Scales-II (OWLS-II), the Peabody Individual Achievement Test-Revised-Normative Update (PIAT-R/NU), and the Test of Adolescent and Adult Language-IV (TOAL-IV). Although these are highly regarded and well-researched tools, even these instruments fail to provide standardization samples of individuals older than their mid-twenties (Penner-Williams, Smith, & Gartin, 2009). A brief description of these formal assessment instruments is included in Table 1.

Of the more popular writing assessment instruments, two current instruments may be considered most adequate for use with the adult population because of their greater age ranges. These include the Woodcock Johnson Tests of Achievement III (WJTA-III) and the Wechsler Individual Achievement Test-III (WIAT-III). While each of these instruments includes normative samples for older individuals, the task requirements and approaches to assessing writing ability are fundamentally different.
Table 1

Summary of Standardized Written Expression Instruments for Adult Populations

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Age Ranges</th>
<th>Description of Writing Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>KTEA-II</td>
<td>4:6-25</td>
<td>This direct assessment of writing ability contains developmentally appropriate story starters in which older individuals write sentences and an essay. Writing samples are evaluated upon structure, content, and sense of audience, and planning. A separate subtest measures spelling ability (Kaufman &amp; Kaufman, 2004).</td>
</tr>
<tr>
<td>OWLS-2</td>
<td>3-21</td>
<td>This instrument provides a sampling of both indirect and direct assessment. The written expression scale of this instrument contains a diversity of writing tasks that includes sentence comprehension and paragraph writing (some tasks are open-ended, while others are more structured). Writing is evaluated upon conventions (spelling, punctuation, etc.) and structural components such as organization, details, and cohesion (Carrow-Woolfolk, 2011).</td>
</tr>
<tr>
<td>PIAT-R/NU</td>
<td>5-22:11</td>
<td>This direct measure of writing ability requires the examinee to compose a writing sample in response to a pictorial stimulus. Written samples are evaluated in terms of organization, grammar, and idea development. Spelling ability is assessed within a separate activity (Markwardt, 1997).</td>
</tr>
<tr>
<td>TOAL-4</td>
<td>12-24:11</td>
<td>This is an indirect measure of written language. Activities include verbal tasks and written response formats. Although this measure includes a sentence combining activity and an editing task, there is no opportunity for the student to construct a written narrative (Hammill, Brown, Larsen, Wiederholt, 2007).</td>
</tr>
<tr>
<td>WJTA-III</td>
<td>2-90</td>
<td>This battery includes indirect measures of written expression ability. Performance in various activities such as spelling, sentence combining, and editing are utilized to derive an overall writing score. There is no opportunity for the student to compose a paragraph or multiple sentence response in this battery (McGrew &amp; Woodcock, 2001).</td>
</tr>
<tr>
<td>WIAT-III</td>
<td>4-50:11</td>
<td>In addition to evaluations of sentence combining and spelling activities, this instrument includes a direct measure of writing skill that is evaluated upon word count, theme development, organization, and grammar (Wechsler, 2009).</td>
</tr>
</tbody>
</table>
The WIAT-III includes an adequate sampling of persons up to age 80 years and written expression is measured through the subtests of Spelling, Writing Fluency, Writing Samples, Editing, and Capitalization/Punctuation. (McGrew & Woodcock, 2001) However, these subtests are used in an indirect assessment of writing skill. These activities do not provide a sample of narrative text of adequate length to evaluate cohesion, organization, planning, sense of audience, or theme development. Arguably, the WIAT-III is an excellent measurement tool when the practitioner wishes to evaluate writing ability through indirect methods. To evaluate writing conventions such as cohesion, organization, planning, etc., a direct measure of writing is required. The only legitimate direct writing assessment instrument for use with adult populations has suffered as a result of its latest revision. The WIAT-III fails to provide the same clinical utility in adult populations as was provided by the WIAT-II. Let us first review the predecessor of the WIAT-III, as the 2nd edition of this instrument may be considered a sound method for directly assessing the writing ability of individuals in college settings.

**WIAT-II**

The Wechsler Individual Achievement Test-2nd ed. (WIAT-II) is a comprehensive and nationally standardized instrument for the measurement of academic skills in children, college students, and adults (Wechsler, 2002). This edition of the Wechsler Individual Achievement Test was published in 2001 and has since been replaced by the more recent 3rd edition. The WIAT-II demonstrates good reliability and validity for interpretation of results. The normative sample consists of 5,586 individuals and reflects the U.S. population in terms of multiple demographic traits (age, sex, geographic region, race/ethnicity, self-education level [for adult sample]) at time of publication. In addition to the school-aged normative group, publishers conducted standardizations with two additional groups in order to create adult norms and expand upon the applicability of this instrument towards older age populations. For the adult norms, two separate samples consisted of a college group and an adult sample. The college group included 707 individuals from both 2-year and 4-year institutions. The adult group includes 500 participants aged 17-85 years. For the adult sample, five distinct age bands were created comprised of 100 members each. These smaller age groups are reflected in the age-based normative tables and include 17-20 years; 21-25 years; 26-35 years; 36-50 years; and 51-85 years. Data for these adult samples were collected during the 1999-2000 and 2000-2001 school years. Reliability of the Written Expression subtest was established by means of test-retest coefficients and intraclass correlations. Multiple approaches to establishing test validity include content related, construct related, and criterion related methods to ensure that items are adequately evaluating the skills they were designed to measure (Wechsler, 2002).

The WIAT-II provides a direct assessment method for writing ability coupled with similar tasks that are included in the WJTA-III. There is a sentence combining activity such as is found on the WJTA-III as well as objective measurements of spelling and word fluency. However, the WIAT-II also includes a direct assessment of the individual’s ability to compose written text in a persuasive format. For practitioners wishing to utilize this type of assessment method in evaluating the writing of adult populations, the WIAT-II is a legitimate assessment instrument supported by numerous empirical studies (Konold & Canivez, 2010; Mayes & Calhoun, 2008; Mayes, Calhoun, & Lane, 2005; Wechsler, 2002). However, the WIAT-II normative data is now outdated and is no longer applicable for making assessment decisions when population comparisons are desired.

**WIAT-III**

In its latest edition, the Wechsler Individual Achievement Test 3rd edition (WIAT-III) has undergone multiple changes in administration and organization as compared to the previous WIAT-II. Rather than simply conducting new normative studies on the previous instrument, test developers restructured the instrument for the purposes of greater utility and applicability for decision-making (Wechsler, 2009). Revisions to the written expression cluster were enacted to offer greater coverage of the various levels of writing ability. These alterations include focused measures beginning with basic processes such as spelling and ranging towards higher order skills such as grammar, mechanics, and paragraph organization. Structural changes in the written expression cluster include the addition of a sentence-building activity along with the previous sentence-combining subtest. Additionally, the essay composition scoring rubrics have been redeveloped for greater rater reliability and there is only one story prompt rather than multiple op-
tions for story starters according to age or grade level. There are three specific scoring rubrics for the written essay, which include word count, theme development/organization, and grammar/mechanics.

The WIAT-III is a structurally and psychometrically sound instrument that has been validated through multiple phases of development (Wechsler, 2009). The normative group provides a representative sampling upon multiple demographic features of the U.S. population that include sex, age, race, grade, geographic location, and parent education levels. The sample group consists of 2,775 individuals for grade-based norms and an overlapping 1,826 individuals for age-based norms. This initial normative sample included grades PK-12 and an age range of 4 years-19:11 years. Evidence of reliability in the written expression cluster was evaluated through test-retest stability. As with the previous edition, internal structure validity and content validity were well established during development of this instrument (Wechsler, 2009).

The WIAT-III normative update containing adult standardization information was released approximately one year after the instrument’s publication (Wechsler, 2010). The adult normative sample included 225 individuals aged 20-50 years. This adult group was divided evenly into three age groupings of 20-25 years, 26-35 years, and 36-50 years. Although the publishers completed an adult standardization, grade based normative information was not included in the WIAT-III. In fact, the interpretive utility of the WIAT-III for adult writing evaluation is limited by the exclusion of these grade-based norms. The WIAT-III adult norms only include age-based information and true comparisons based upon grade-based norms, age and grade equivalents, or growth scale values are not included. According to the manual, estimates of these grade-based scores may be obtained by comparing the adult’s score results to the school age norm tables. The practitioner may either refer to PK-12 charts or elect to change the adult examinee’s age to less than 20 years if using the computer scoring software. Nevertheless, these estimations are imprecise since the PK-12 normative sample is not intended for use with adult populations. Additionally, the use of the PK-12 sample is only applicable when the adult examinee’s level of functioning is below grade 12. Unlike the WIAT-II, there are no college based normative samples included in the WIAT-III. As a result, the previous option of comparing the education level of adults that was offered by the WIAT-II is no longer a possibility with the WIAT-III. The clinical utility of the WIAT-III for use in postsecondary settings has been severely lessened as a result of failure to include college based normative samples and the omission of age and content appropriate essay prompts.

The Revised Normative Groups

Comparisons between the adult normative groups that were espoused for the 2nd and 3rd editions of the WIAT result in striking discrepancies. First, the WIAT-II included an adult normative group as well as a college normative group. The adult group included 500 persons and the college sample included 707 persons (Wechsler, 2002).

The WIAT-III only includes an adult normative group of 225 individuals, which is about half as large as that included in the 2nd edition (Wechsler, 2010). The age range of the WIAT-III terminates at 50 years, while the 2nd edition included an age range up to 85 years. Likewise, there is no inclusion of a college normative group in the WIAT-III. As a result, the WIAT-III not only lacks information for college grade level comparisons, but also the scope of the sample included in the WIAT-III is significantly less than what appeared in the 2nd edition.

The Wechsler Story Prompts

Previous research has demonstrated the importance of the story starter or cue in writing assessment (Cole et al., 1997; Gregg et al., 2008; Hooper et al., 1994; Muenz et al., 2000). Although the WIAT-III may provide greater utility for the practitioner, revisions to the written expression cluster do not provide the optional story starters that were included in the WIAT-II. For practitioners wishing to evaluate an adult writer, the WIAT-III story prompt may appear elementary or inappropriate. Rather than conducting standardization studies for adults using an alternate story prompt, the adult normative group completed the “favorite game” writing activity. This story prompt is in stark contrast to the previous prompt included in the WIAT-II, which required the adult writer to complete a persuasive essay arguing either for or against free tuition in higher education settings or alternatively writing a persuasive essay stating an opinion concerning the adoption of daylight savings time.

The writing tasks included within the WIAT-II and WIAT-III are markedly different. The WIAT-II stories
are persuasive in nature and the WIAT-III story task appears to be that of a more expository or narrative format. While the lack of a persuasive writing requirement in the WIAT-III is not detrimental in and of itself, the use of a writing prompt that may be developmentally inappropriate for adults and the omission of an alternative writing prompt is concerning.

**Conclusion**

Research into the assessment of written expression, as compared to other academic areas, may still be a burgeoning field of study. However, despite growth in this area of psychoeducational assessment in recent years, the adult population has been largely overlooked by test developers. There are few options for the practitioner wishing to employ standardized measurement for the written expression abilities of adults. This may be expected and is certainly defensible from a pragmatic stance, since the collection of normative data for these populations may not be a financially beneficial undertaking for the widely marketed publishing groups. Arguably, the only direct writing assessment instrument truly validated for adult populations, the WIAT-II, is now outdated with the recent revisions in its 3rd edition. Although the newest edition of the WIAT-III includes adult sample groups, this instrument suffers from two significant limitations. Unlike the previous edition, the story cue is the same for all age groups. The “favorite game” writing activity is inappropriate for use with adult populations. Additionally, the WIAT-III adult norms are significantly lacking when compared to those that were included in the WIAT-II. While there was an adult sample included for the WIAT-III, it pales in comparison to the extensive sampling provided for college and adult norms that benefitted the WIAT-II.

These two shortcomings in this newest edition significantly weaken the usefulness of the WIAT-III in postsecondary school environments, which may be the arena in which measurements of adult writing are most likely to occur. Due to the limitations of the WIAT-III, there is no longer a viable standardized direct writing assessment for use with college populations. Despite the option to measure the written expression abilities of college students indirectly with the WJTA-III, there is now a real need for a direct writing assessment instrument for this population. For those few practitioners whose primary clientele include college-aged students and adults, the measurement of written expression becomes a complicated and often ambiguous undertaking. Future focus upon test development for adult writers would greatly benefit these practitioners, especially those working within postsecondary settings. The WIAT-III appears to be a highly useful instrument, but the lack of appropriate writing tasks and adult normative samples are in stark contrast to those provided in the previous edition. The lackluster attempts at a college and adult writing assessment measure contained in this newest edition of the WIAT-III is astoundingly limited. This oversight gives a sense that the adult writer was but a mere afterthought by test developers.

This discussion of written expression assessment may not be salient to those disability service providers who do not directly assess their students for formal diagnostic purposes. However, each of the previously discussed instruments is viable for determining functional impairments and guiding the determination of accommodations or strategy instruction in written expression. Readers who possess the appropriate training in the administration of these achievement tests may utilize them to gauge academic functioning. Although diagnostic decisions are reserved for those professionals who are otherwise qualified, most readers will be able to make use of these instruments to determine functional impairments.

While the purpose of this article has been to highlight the need for a better diagnostic tool for disorders of written expression in adult populations, all readers should be aware of the current lack of adequate direct writing assessment instrumentation. For those readers who review disability documentation or psychological reports, knowledge of the form of written expression assessment (i.e., direct or indirect) and consideration of the task demands of instruments can influence decision making in regards to academic supports. Some writing tasks on popular instruments are not reflective of expectations in postsecondary settings. For example, the writing activities on the WIAT-II and WIAT-III are markedly different. Although formalized assessment of writing is often necessary to determine the presence of a learning disorder, the determination of functional impairments in writing may be assessed through less structured methods than those included on the previously discussed standardized writing assessments.

The authors believe that making decisions about functional impairments after a student is determined eligible for disability services may be accomplished through a variety of techniques. Consider reviewing classroom samples or school records to determine areas
of difficulty. Also consider scores from the Scholastic Aptitude Test (SAT) as a source of information. Some individuals may wish to obtain an informal writing sample from the student. An informal but highly valuable option would be to ask students to write a brief persuasive essay about a topic of your choosing. Give the student a limited amount of time to compose the essay. Then you may review the student’s writing to inform your decision-making about possible accommodations or strategies for improvement. If the student does not plan effectively and the essay lacks acceptable structure, then writing workshops or campus writing services may be appropriate. Basic instruction in paragraph construction and planning may have a tremendous impact. If the student experiences difficulties due to poor handwriting, a word processor could be an appropriate accommodation. The need for spellcheckers is also quickly evident if you take the time to review an actual sample of writing from a poor speller. A student’s failure to complete a writing assignment within specified time limits may warrant use of extended time.

Assessment practices for measuring written expression are varied, with no definitive consensus amongst researchers and experts on the issue. Although the authors are not fond of the newest changes that have occurred with the WIAT-III written expression subtest, we remain adamant in our opinion that direct samples of writing are invaluable in determining intervention strategies for improvement. While it is true that there are far fewer options for written expression assessment in older populations of students, this is in no way reflective of a decreased importance of this life skill as one matures. The procedures for determining the presence of written expression disorders may have been impacted by the latest revisions to the WIAT-III, but the ability of disability service providers to determine functional impairments remains as dynamic a process as ever. Be creative and deploy holistic techniques to examine direct samples of writing or choose quantitative measurement techniques by evaluating writing indirectly. Research and focus upon the written expression of adult populations is obviously limited. Disability service providers enjoy a degree of flexibility when using professional judgment to make accommodations decisions. In the case of written expression, that flexibility is not only afforded by your competency as a professional but also as an unfortunate necessity due to the decreased availability of standardized written expression assessments for college students.

References


Individuals with Disabilities Education Improvement Act of 2004. 20 USC 1400


Ucgun, D. (2011). The study on the writing anxiety levels of primary school 6, 7, and 8th year students in terms of several variables. *Educational Research and Reviews, 6*(7), 542-547.


### About the Authors

Daniel McNair, a native of Sandersville, GA, received a BA degree in Communication Arts from Georgia Southern University in Statesboro, GA. Graduate work includes Master’s and Educational Specialist degrees in School Psychology from the same institution. He completed school psychology practicum and internship requirements within Bulloch County Public Schools. Mr. McNair is a certified school psychologist in the state of Georgia and is currently employed on the campus of his alma mater with the Regent’s Center for Learning Disorders- one of three such centers in the state- that specializes in psychoeducational evaluations of postsecondary students. His research interests include written expression assessment, cyberbullying, anxiety disorders, social deviance-particularly as it applies to younger populations, and cultural hegemony. He may be reached at: danielmcnair@georgiasouthern.edu

Dr. Toi Curry, a native of Yonkers, NY, earned her Bachelor’s degree in psychology from Marist College. She then went on to earn her Master’s and Doctoral degrees in Clinical Psychology from The Institute for Clinical Psychology at Widener University. While in her doctoral program she specialized in Clinical Neuropsychology and School Psychology. She is a certified School Psychologist in the state of Pennsylvania and holds licenses to practice psychology in both New York and Georgia. Dr. Curry completed a 2-year postdoctoral fellowship in Clinical Neuropsychology at the Lenox Hill Hospital’s Center for Attention and Learning Disorders in New York City. Research interests include reading and math disabilities across the lifespan, language development, recovery and reintegration into education following head injury, and the impact of pregnancy complications on the developing brain.