Issues in the Identification of Learning Disabilities

Yes, Virginia, There is a Severe Discrepancy Clause, but is it too much ado about something? - This article takes the view that severe discrepancy is an often misunderstood and misused concept. The article is posted on the APA Division 16 website and can be downloaded as an Adobe Acrobat Reader file (PDF).

Links to Learning Strategies - A list of hyperlinks leading to effective teaching and learning strategies

- Effects of speed (or lack thereof) - What would be the effects on a child's WISC-III scores if that child did exactly what we often tell kids to do "Do it slow and get it right"? You might be surprised.

- LD: Major Papers - In preparation for the upcoming reauthorization of IDEA, the US Office of Special Education Programs (OSEP) has commissioned leading researchers in the field of Learning Disabilities to write major papers on nine topics. We (John and Ron) must admit that we can't agree with them all, but they make for an interesting read.

- IF only this were made up - The LD definitions and facts on this page really belong on our humor page, except they are real!

Learning Strategies Guides - Links to study guides that are available on-line. Many of these are maintained by the academic services centers at colleges and universities, but much of their information is useful for high school and middle school students, as well. (Our thanks to Jonas Taub for allowing us to post this)

WISC-III Venn Diagram - "It came to me in the car between Hackensack and Portsmouth. I realized that it was what I have been doing with Wechsler scales back to around 1939 or whenever I started with them, so I made the first draft tracing my cognac glass on a blank transparency with an overhead marker in the motel room. The ink made the cognac taste awful."

Reading Programs - Just a few good web sites for information regarding reading programs
Attention in School: Some Advice - Steve Edelman has written a page that provides ways to increase a child’s ability to pay attention and concentrate in school without having him or her take medication.

Bender Gestalt and Emotional Indicators - A demonstration of the need for caution when using the, so called, "emotional indicators" of this popular visual motor task.

Age Vs Grade Based Scores - This is an archive of a discussion carried on the NASP List Serve.

Example of Reading Comprehension Formulae - This is an example of how slight changes in a sentence may have great impact on the readability level.

Basic Disorders - Comments on the issue of "disorders in basic psychological processing" and examples of such disorders.

Evaluation Mistakes - A list of common errors made by evaluators. Good to know so you don't make the same mistakes.

Deficits - A description of common behavior indicators in 10 domain areas.

EBD - Everybody's Disabled. An essay on the proposed change in education labeling that would change the Seriously Emotionally Disturbed (SED) category into Emotional/Behavioral Disturbance (EBD).

Learning Disabilities Talking Points - 30+ issues relating to the identification of learning disabilities in children.

Double Standards in Testing or Do We Need a Gun Control Law for Testers - Commentary on the state of evaluation in the field of learning disabilities.

Suspension Notes - Random notes and cautions regarding suspension of students.

Post-Otitis Auditory Dysfunction (POAD) - A serious and underestimated cause of multiple learning problems in school children. Despite Boucher's (1986) dramatic findings, the problems of POAD often go unrecognized and untreated.

Grade Equivalents - The OAT CEREAL Achievement Test as an example for how the use of grade equivalents is a silly idea.

Logical Steps in the Determination of a Learning Disability - Thank you to John Willis for providing a synopsis of the steps in LD evaluation.

Predicting Achievement from IQ - How to determine achievement based on the IQ score chosen and the correlation between IQ and
Learning Disabilities: A very brief explanation.

**Severe Discrepancy** How to determine severe discrepancy using a method described by Cecil Reynolds.

**New Severe Discrepancy 2** How to determine severe discrepancy using a method described by Hubert Lovett. Go to bottom of this page to download a template. Thank you Hubert for all the work.

This section expands upon the issues presented in Predicting Achievement from IQ

**Subtest order and standardization rules** Does the order of administration matter when it comes to test results? Read this lively discussion about that very issue (featuring Ron Dumont; Laurie Ford; Richard Woodcock; Mark Daniel; Samuel O. Ortiz; Kevin McGrew; John Willis, Catherine A. Fiorello)

**ADHD and the Mesulam** Can a simple visual scanning task help in the diagnosis of ADHD. This article describes the norming of the Mesulam with children and then a comparison of those norms to a sample of ADHD children.

**Test Behaviors and Testing the Limits** Some hopefully useful suggestions about test behaviors seen during testing as well as some ideas on how to "test the limits."

**Not all tests scores are alike** Just because you gave a test doesn't mean the results are easy to understand. Explanations of some testing concepts.

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Yes, Virginia, there is a Severe Discrepancy Clause, but is it Too Much Ado About Something?

Ron Dumont, Fairleigh Dickinson University  
John Willis, Rivier College  
Guy McBride, Burke County Public Schools, NC.

There are three kinds of people in this world - those who can count and those who can’t.  
Unknown

Even if you are on the right track, you’ll get run over if you just sit there.  
Will Rogers

If you’re enjoying this program just half as much as we’ve enjoyed doing it, then we’ve enjoyed it twice as much as you.  
Michael Palin

Bad facts make bad law.  
Frank Zappa

A respect for the law of parsimony is a characteristic of science, but educational psychology’s penchant for simple answers to questions of complex behavior, particularly in the area of learning disability, has tended toward paucity rather than parsimony of explanation.  
John McLeod

Dear Editor of The School Psychologist:  
I am 28 years old. Some of my friends say there is no need for a numerical severe discrepancy when determining a specific learning disability. Papa says, "If you see it in The Newsletter, it’s so." Please tell me the truth, do we need a severe discrepancy to identify a specific learning disability?  
-Virginia

The answer to this question seems to depend upon whether you are asking about the presence of a learning disability or the presence of an educationally disabling condition. Contrary to popular opinion, Virginia, there is no federal requirement for any sort of mathematical measurement of "severe discrepancy," there is no federal requirement to base any such comparison on test scores at all, and a severe discrepancy may, in fact, be completely irrelevant to the determination of a specific learning disability (SLD).

The federal criteria carry the force of law and virtually every decision and letter we have reviewed speaks to the need for an Individualized Education Plan (IEP) team to determine that there was a severe discrepancy. There are explicit dicta within the federal definition and the regulations that the disability reflect a disorder in psychological processing; and that the cause of the child's deficiencies not be primarily the result of a visual, hearing or motor impairment; mental retardation; or social, economic, or environmental factors CFR sections 300.7 (10); 300.541. Additionally, CFR section 300.534 prohibits the classification of a child under any classification if:

(1) The determinant factor for that eligibility determination is—
   (i) Lack of instruction in reading or math; or (ii) Limited English proficiency; and
   (2) The child does not otherwise meet the eligibility criteria under Sec. 300.7(a).

Probably nothing related to the identification of specific learning disabilities causes multidisciplinary teams (MDTs) more problems than the decisions they are asked to make regarding whether the child has or does not have a severe discrepancy. Teams are often struggling for clear guidelines to help objectify the process. Questions related to the area of severe discrepancy often include: Does the Individuals with Disabilities Education Act (IDEA) actually demand that a team find a severe discrepancy? Is a severe discrepancy defined or clarified anywhere in IDEA?

Continued on page 4
Professional Communities of the Future

Jack Cummings, Indiana University

Professional organizations give members a sense of community. As members of Division 16, we share values of caring and working for the benefit of children. We share knowledge of tools for academic and social/behavioral interventions designed to improve learning environments. We strive to understand children’s difficulties in more than a within-child framework, to consider classroom dynamics, and reflect on broader systems level issues. In this context we influence schools as well as association of School Psychologists, but not the British Psychological Association. The travel time and costs prohibit me from regular attendance at the British conferences.

Psychologists across various regional and national psychological associations encounter children who have learning difficulties, as well as those who struggle with behavior in classrooms and those with problems initiating and maintaining social relations with peers. These problems are not place-based, they are common difficulties of children in schools regardless of region or nation. A recent email on the EPNET, a listserv that serves educational psychologists in Great Britain, illustrates the point. The psychologist requested information on ways to address the problem of bullying. In response, individuals on the listserv suggested various web sites that address the problem.

Most members who have been part of a listserv for a number of months will note that these electronic venues share characteristics with conventional face-to-face communities. Norms develop for what are socially accepted communications. If someone strays from the norm, other members will remind the individual to behave properly. Over time, a shared history of interactions develops. A sense of identity develops and members share a form of community where they are able to exchange or share ideas and resources.

Listservs offer a glimpse of potential future electronic communities. Recently developed Web-based software allows a presenter to communicate with a small or large group. The powerful value-added aspect of using the Web is that the location of the audience is irrelevant. The meeting can be open to anyone with access to the Internet, or it can be closed by requiring the user to enter the conference space with a password. A Web-based presentation can be quite similar to those conducted in a face-to-face format. The speaker may present a sequence of PowerPoint slides, be heard (audio only) or be heard and seen in real-time video. The audience can submit questions or be polled for their reactions to points made by the speaker. Software exists to allow sharing of a "whiteboard" space, such that the presenter and the participants can work together as they would on a conventional blackboard.

Other web tools exist to allow asynchronous communications. Web forums have the advantage of requiring neither geographic proximity nor assembling at the same time, hence the term asynchronous. Within the Indiana University School Psychology Program we have used web forums for debates.

Continued on page 15

Editor's Message

Vincent C. Alfonso, Fordham University

As I begin my first of three years as editor of The School Psychologist, I thought it would be a good idea to make some opening remarks. First, I’d like to recognize Dr. Steven G. Little for his significant contributions to the newsletter these past six years. Steve has edited TSP with great professionalism and scholarship. He has helped me tremendously in my transition to editor. Second, I’d like to bring your attention to the editorial staff of TSP by directing you to the back cover. A perusal of the names and institutions will, in all likelihood, impress you. I am grateful to Dr. Linda A. Reddy, associate editor, for taking a leadership role in the production of the newsletter. Third, I invite you to be a part of the newsletter by contributing materials such as an article or test/book review. Although there are printing deadlines, I can assure that most submitted material will be printed in an issue. Fourth, I invite you to submit comments to Linda (reddy@alpha.fdu.edu) regarding any printed information in the newsletter especially the feature article. Finally, Dr. Angeleque Akin-Little has agreed to edit the "people and places" column that was a part of previous issues. Angeleque can be reached at psyaza@hofstra.edu. I hope that you enjoy reading TSP during the next three years!
Division 16 Executive Committee

President
Jack Cummings, Ph.D.
Education #4038
Indiana University
201 North Rose Ave.
Bloomington, IN 47405-1006
(812) 856-8327
cummings@indiana.edu

Vice President
Deborah Tharinger, Ph.D.
SZB 504
University of Texas
Austin, TX 78712
(512) 471-4407
(512) 475-7641 (fax)
dtharinger@mail.utexas.edu

Past President
Jack Cummings, Ph.D.
Education #4038
Indiana University
201 North Rose Ave.
Bloomington, IN 47405-1006
(812) 856-8327
cummings@indiana.edu

Secretary
Michelle Schicke Athanasiou, Ph.D.
Div. of Professional Psychology
University of Northern Colorado
Greeley, CO 80639
(970) 351-2356
(970) 351-3626 (fax)
mathan@edtech.unco.edu

Treasurer
Sharon A. Missiaen, Ph.D.
5345 Canvasback Road
Blaine, WA 98230
(360) 371-7836
(360) 371-0527 (fax)
shindes@gte.net

Vice President of Professional Affairs
Colette L. Ingraham, Ph.D.
Department of Counseling and School Psychology
MC 1179, College of Education
San Diego State University
San Diego, CA 92182-1179
(619) 594-8605
(619) 594-7025 (fax)
ingraham@mail.sdsu.edu

Vice President of Educational, Training, & Scientific Affairs
Alberto Bursztyn, Ph.D.
Brooklyn College - CUNY
2900 Bedford Avenue
Room 2107, James Hall
Brooklyn, NY 11210
(718) 951-5214
(718) 951-4816 (fax)
aburstyn@brooklyn.cuny.edu

SASP Representative
Matt Turner
1559 Chase Arbor Common
Virginia Beach, VA 23462
(757) 467-2378

Vice President of Social and Ethical Responsibility & Ethnic Minority Affairs
Robyn S. Hess, Ph.D.
School of Education, C.B. 106
P.O. Box 173364
Denver, CO 80217
(303) 556-6784
Robyn.Hess@co.cudenver.edu

Council Representatives
Jonathan Sandoval, Ph.D.
Division of Education
One Shields Avenue
U.C. Davis
Davis, CA 95616-8579
(916) 752-3198
(916) 752-5411 (fax)
jsandoval@ucdavis.edu

Historian
Thomas K. Fagan, Ph.D.
Department of Psychology
The University of Memphis
Memphis, TN 38152
(901) 678-4676
tom-fagan@mail.pscy.memphis.edu

Editor, School Psychology Quarterly
Terry B. Gutkin, Ph.D.
117 Bancroft Hall
Dept. of Educational Psychology
University of Nebraska-Lincoln
Lincoln, NE 68588
(402) 472-8317
tgutkin1@unl.edu

The School Psychologist is published four times per year by the Division of School Psychology (Div. 16) of the American Psychological Association. Subscriptions are free to members of the Division. For information about subscription rates, submission of articles or advertising write: Vincent C. Alfonso, Ph.D., Fordham University, Graduate School of Education, 113 West 60th St., New York, NY 10023.

The Division reserves the right to edit all copy and to refuse ads that are not in consonance with its principles. The publication of any advertisement by the Newsletter is not an endorsement of the advertiser nor the products or services advertised. Division 16 or APA is not responsible for any claims made in the advertisement. Advertisers may not, without prior written permission, state in any subsequent advertisement the fact that a product or service has been advertised in a Division 16 or APA publication.

Division 16 Home Page
Check out the Division 16 home page at:
http://www.indiana.edu/~div16/

PUBLIC HATION SCHEDULE
Issue Month/No. Closing Date for Submission of Materials Printing Date Mailing Date*
Winter (1) December 1 December 15 January 15
Spring (2) January 15 March 15
Summer (3) February 1 May 15 July 15
Fall (4) March 15 October 15

* Allow 3 to 6 weeks for delivery of 3rd class mail.
How severe is severe? What is the severe discrepancy discrepant from? What IQ or other ability score (e.g., Wechsler Full Scale, Verbal, Performance IQ; Differential Ability Scales General Conceptual Ability or Cluster Score, Woodcock-Johnson Tests of Cognitive Ability III General Intellectual Ability or Cluster Score, etc.) can be used when determining IQ-Achievement discrepancies? Are there other, acceptable ways to determine severe discrepancy besides using the IQ-Achievement test score comparison?

The real problem is that neither the federal definition nor the federal regulations have ever defined any of the key terms (e.g., "achievement," "intelligence," "severe," and even "primarily"). Definitions of these terms are left to the IEP teams with whatever guidance their states provide them. In some instances, States have provided guidance for what these terms might mean, but when IEP teams have taken them too literally, the Office for Special Education and Rehabilitative Services (OSERS) has slapped their hands, with the injunction that "no one formula" may be used to establish eligibility.

Various professions have encouraged different definitions of learning disabilities. These definitions, although often very similar to each other, have enough difference between them to cause problems for MDIs when it comes to choosing among them. Below are a few of the many conceptualizations of what constitutes a learning disability. [We have included one State definition (New York) as an example of how the State definitions may differ from other definitions.]

IDEA [§300.7 (c)(10)]

Specific learning disability . . . means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in an imperfect ability to listen, speak, read, write, spell, or to do mathematical calculations. The term includes such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. . . . The term does not include learning problems that are primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

The Learning Disabilities Association of America (LDA)

Specific Learning Disabilities is a chronic condition of presumed neurological origin which selectively interferes with the development, integration, and/or demonstration of verbal and/or nonverbal abilities. Specific Learning Disabilities exist as a distinct handicapping condition and varies in its manifestations and in degree of severity. Throughout life, the condition can affect self-esteem, education, vocation, socialization, and/or daily living activities. [Association for Children with Learning Disabilities (1986). ACLD Description: Specific Learning Disabilities. ACLD Newsbriefs, Sept./Oct. (166), 15. Note: The Association for Children with Learning Disabilities is now the Learning Disabilities Association of America.]

The National Joint Committee on Learning Disabilities

Learning disabilities is a general term that refers to a heterogeneous group of disorders manifested by significant difficulties in the acquisition and use of listening, speaking, reading, writing, reasoning, or mathematical abilities. These disorders are intrinsic to the individual, presumed to be due to central nervous system dysfunction, and may occur across the life span. Problems in self-regulatory behaviors, social perception, and social interaction may exist with learning disabilities but do not by themselves constitute a learning disability. Although learning disabilities may occur concomitantly with other disabilities (e.g., sensory impairment, mental retardation, serious emotional disturbance), or with extrinsic influences (such as cultural differences, insufficient or inappropriate instruction), they are not the result of those conditions or influences. [National Joint Committee on Learning Disabilities (1990).]


315.00 Reading Disorder

A. Reading achievement, as measured by individually administered standardized tests of reading accuracy and comprehension, is substantially below that expected given the person's chronological age, measured intelligence, and age-appropriate education.

B. The disturbance in Criterion A significantly interferes with academic achievement or activities of daily living that require reading skills.

C. If a sensory deficit is present, the reading difficulties are in excess of those usually associated with it.

[Note: definitions of Mathematics Disorder (315.1), Disorder of Written Expression (315.2), and Learning Disorder Not Otherwise Specified (315.9) are similar.]

Additionally:

Learning Disorders are diagnosed when the individual's achievement on individually administered, standardized tests in reading, mathematics, or written expression is substantially below that expected for age, schooling, and level of intelligence. The learning problems significantly interfere with academic achievement or activities of daily living that require reading, mathematical, or writing skills. A variety of statistical approaches can be used to establish that a discrepancy is significant. Substantially below [italics in original] is usually defined as a discrepancy of more than 2 standard deviations between achievement and IQ. A smaller dis-
Yes, Virginia, there is a Severe Discrepancy Clause, but is it Too Much Ado About Something?

crepancy between achievement and IQ (i.e., between 1 and 2 standard deviations) is sometimes used, especially in cases where an individual's performance on an IQ test may have been compromised by an associated disorder in cognitive processing, a comorbid mental disorder or general medical condition, or the individual's ethnic or cultural background. If a sensory deficit is present, the learning difficulties must be in excess of those usually associated with the deficit...Individualized testing is always required to make the diagnosis of a Learning Disorder...Learning Disorders must be differentiated from normal variations in academic attainment [boldface in original] and from scholastic difficulties due to lack of opportunity, poor teaching, or cultural factors (pp. 49-51).

New York [Part 200.1, (mm) (6)]

"Learning disability" means a student with a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which manifests itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations. The term includes such conditions as perceptual handicaps, brain injury, neurological impairment, minimal brain dysfunction, dyslexia, and developmental aphasia. The term does not include students who have learning problems which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage. A student who exhibits a discrepancy of 50 percent or more between expected achievement and actual achievement determined on an individual basis shall be deemed to have a learning disability.

While all of these definitions have merit, none has the force of law in establishing eligibility under IDEA, which provides legal entitlements to children meeting the congressional and regulatory criteria. State definitions and regulations cannot supersede the federal criteria although they can exceed them—that is, they can provide a higher level of entitlements, but they cannot restrict the entitlements provided under the act.

In 1997, Congress passed a substantial revision of IDEA. In 1999, OSERS published its final regulations implementing IDEA, defining a specific learning disability in the same way as it did in 1977. It is interesting to note that in the Federal definition of a learning disability [§300.7 (c)(10)] as well as in the definitions of most advocacy groups and States, there is no mention of the term "severe discrepancy." One must read further in the relevant Regulations in order to find any mention of severe discrepancy. In IDEA the term is found at §300.541 under the heading "Criteria for determining the existence of a specific learning disability" (below, italics added):

(a) A team may determine that a child has a specific learning disability if
(1) The child does not achieve commensurate with his or her age and ability levels in one or more of the areas listed in paragraph (a)(2) of this section, if provided with learning experiences appropriate for the child's age and ability levels; and
(2) The team finds that a child has a severe discrepancy between achievement and intellectual ability in one or more of the following areas:
   (i) Oral expression.
   (ii) Listening comprehension.
   (iii) Written expression.
   (iv) Basic reading skill.
   (v) Reading comprehension.
   (vi) Mathematics calculation.
   (vii) Mathematics reasoning.

We would note that, in this one instance, the Regulations say a team "may determine." Is there a difference between "may determine" and "must determine" or "this is the only way to determine?" Note that the Regulations do not define a severe discrepancy.

There appears to be a distinction between the definition of a learning disability and the criteria used to establish qualification for special education services based upon a disability. The definition of specific learning disability establishes the presence of a "disorder" which is manifesting itself in one or more areas of academic achievement as listed above. With respect to other disabilities, the term "severe" and the phrase "adversely affects educational performance" are generally thought of as referring to some absolute deficit. However, when considering SLD eligibility, "severity" must be measured against the child's expected performance, not against some arbitrary general standard. The common requirement is that in determining eligibility under the SLD classification as in others, the team must also find that the child needs special education in order to receive a free appropriate public education (FAPE)—still another term in special education law that is not operationally defined. [Although it is generally accepted by the courts that the benefits expected must be non trivial or meaningful, "need" must be determined by looking at the whole child—a team may not assume that a child does not need special education services simply because he or she is receiving passing grades and/or is being passed from one grade to the next (see Florence County v. Carter, 1991).]

If this IDEA 97 regulation [§300.541] is taken at face value, there must be a severe discrepancy between ability (IQ?) and achievement (test scores?) in order for a student to be identified as having a learning disability. Obviously one should not read this section of the regulations in isolation and believe that the mere presence of a severe discrepancy establishes the presence of a learning disability. The severe discrepancy may be the result of factors other than a "disorder in one or more of the basic psychological processes." As noted above, the Regulations require that the multidisciplinary team rule out other factors before determining eligibility as having a learning disability. IDEA 97 Final Regulations,
Yes, Virginia, there is a Severe Discrepancy Clause, but is it Too Much Ado About Something?

Subpart A §300.7 (c)(10)(ii) notes:

Disorders not included. The term does not include learning problems that are primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

The criteria for determining SLD involve a multi-step process. One step is for the team to determine that the child is not achieving at a level commensurate with age and ability when provided with educational opportunities. This forces the team, in the ideal world, to rule out apedagogia (lack of instruction) and dyspedagogia (inappropriate instruction) as the source of the learning problem. Unfortunately, teams seldom determine, or even consider, that the source of the child's learning problems is the teacher(s) or the administrator(s), or a mismatch between the child and the curriculum. Instead, the child usually is identified as the problem.

Additionally, as a condition of eligibility, the team would need to determine that the child required special educational services in order to receive FAPE [§300.7(a)(1)].

- Okay, I see the difference between the specific definition of a child with a learning disability and the general definition of one who meets the "criteria for eligibility for special education services," but doesn’t the IDEA language, and more specifically the language of the courts, offer more definitive answers than these?

Sorry Virginia, but the courts, in many instances, seem to offer little true guidance in these matters. For example, taking it out of the realm of SLD one would at least expect Circuit Court judges to have some sort of consensus over procedural matters—due process is, after all, their bread and butter. But even in such seemingly "legal" areas of special education law as "Who bears the burden of proof?" the Federal Circuit Courts are split as to how that should be determined. "Is additional evidence allowed upon appeal?" is another seemingly clear-cut legal question upon which the courts should be expected to rule with decisiveness. The First Circuit Court allows supplemental evidence but no embellishments of previous testimony; the 3rd Circuit Court says district courts may exclude additional evidence but with discretion; the 4th says if the evidence was available at the time of the due process hearing, district courts may exclude it; the 6th allows new evidence, but not on new issues; the 7th allows "supplemental" evidence; the 9th agrees with the 1st and 7th; and the 11th Circuit Court threw up its hands and said that whether additional evidence will be allowed is up to the "sound discretion" of the district court judge (TSE, December 15, 2000, p. 5). We make this point only to emphasize how perilous it is to base a conclusion on any single case, much less a due process hearing.

Due process decisions only have force in the system wherein they were issued, and then only if not overturned by a court of competent jurisdiction. There is a danger in over-generalizing the results, because they may contain some "bad law," that is, legal decisions based in part on findings that would have been overturned on appeal, or decisions agreed to by the school simply because settling was more cost effective than litigating over principle. For example, a hearing officer might find that a child had failed to make progress based on standard scores that remained constant over time, certainly a finding that can be appealed—but if the IEP was inappropriate, the hearing officer’s possible error would never be reviewed and, therefore, never be corrected.

Additionally, some caselaw has force only in the region in which the case was decided. For example, a fairly well-known case, Larry P. v. Riles, which sought to prohibit the use of intelligence tests in assessing African-Americans was upheld in The 9th Circuit Court (1984) but it applies only to California. No other circuit has ruled against using IQ tests for African-Americans (OCR, 2000).

- Given what the law says, isn't the use of some statistical formula or computer program not only required, but probably the only way to determine a severe discrepancy?

First, Virginia let us say that discrepancy formulae are statistical nightmares (Aaron, 1997; McLeod, 1974; Reynolds, 1990). Furthermore, the absence of some discrepancy should not be used as the sole criterion to exclude children from a specific learning disability (SLD) classification. Discrepancies might be better thought of as preventive in nature, not exclusive. Since the regulations provide little guidance and no specific requirement regarding any particular way of measuring the discrepancy, teams appear to have great latitude regarding how they make discrepancy decisions.

One state, New York, added a sentence into its definition of Learning Disability (see above) that notes "A child found to have a 50% discrepancy between expected achievement and actual achievement shall be deemed to have a learning disability." Almost 20 years ago, in Riley vs. Ambach (United States Court of Appeals, Second Circuit 668 F.2d 635; 1981 U.S. App. December 16, 1980, Argued May 19, 1981, Decided), the court ruled that New York could not use the absence of the "50% discrepancy" to exclude children. The Court concluded that the use of the 50% standard interferes with the proper identification of learning disabled children since it operetes to eliminate consideration of factors and the use of techniques that do not, "given the present state of the art," lend themselves to quantification. The clause has come to be defined, maybe as it should have been all along, as a presumptive clause; if all things are in place you can presume the child eligible. But, if a numerical discrepancy isn't clearly evident, you cannot, on that fact alone,
exclude the child. We argue that this is true of IDEA. IF there is a discrepancy you MAY identify the child as having a Learning Disability; BUT if there is no numerical discrepancy, you cannot automatically exclude the child. Similar to catching the child red-handed in the act of committing a crime—you have the evidence, presume the child guilty. But what if you don't catch the child red-handed, does that mean the child could not be guilty? Obviously, there may be other convincing evidence.

Second, as far back as August 23, 1977, the newly created OSERS published its first regulations in implementation of Public Law 94-142. Those regulations included the following statement:

“No single procedure is used as the sole criterion for determining an appropriate educational program for a child” (emphasis ours).

One might argue that if a MDT uses a severe discrepancy as the sole criterion to exclude a child, it has violated this clause of IDEA. The Department of Education indicated an intent to review current research on learning disabilities and to make changes to the definition, but as of May, 1999, it had not done so, saying this in the appendix to the regulations:

“[T]he Department plans to carefully review research findings, expert opinion, and practical knowledge over the next several years to determine whether changes should be proposed to the procedures for evaluating children suspected of having specific learning disabilities.”

Many teams and districts are tempted to establish mathematical formulas to identify students with specific learning disabilities. Some commercially produced computer scoring programs automatically calculate discrepancies. These formulas attempt to measure the severity of the discrepancy between ability and achievement. We advise that such formulas not be used as the be-all and end-all of learning disability determination. Over reliance upon a "magic number" can expose a system to an adverse finding (e.g., Long Beach Unified School District, 1998). It is therefore not wise to try to use automatic formulas to establish specific learning disabilities, mental retardation, hearing impairment, or any other educational handicap. If a team or district were absolutely determined to use a "discrepancy formula," professional "best practice" would suggest the following rules (see also McLeod, 1974 and Reynolds, 1990. These issues are also discussed thoughtfully in Flanagan, McGrew, and Ortiz (2000) and Mather and Schrank (2001).]

1. The formula should use some form of standard score (including normal curve equivalents). Percentile ranks cannot be used for mathematical comparisons. Grade and age equivalents should not be used for making eligibility decisions. The ability and achievement tests should, of course, be reported with the same statistics (grade equivalents are deceptively simple and may misrepresent the severity of a child’s problem (e.g., Willis & Dumont, 1998, p. 223). Therefore, they should be used with caution. However, they can sometimes be used more effectively than standard scores in documenting progress (e.g., the 9th Circuit Court in a recent case wrote, "HISD argued that passing marks and advancement from grade to grade were sufficient indicia to satisfy the IDEA). And on this dispute the district court is correct that a disabled child's development should be measured not by his relation to the rest of the class, but rather with respect to the individual student, as declining percentile scores do not necessarily represent a lack of educational benefit, but only a child's inability to maintain the same level of academic progress achieved by his non-disabled peers " (Houston Independent School District v. Caius, 2000).

2. The formula should take into account the phenomenon of "regression toward the mean," although a number of states have adopted non-regression formulas for the sake of simplicity. When one attempts to predict one variable from another, the predicted score is likely to be closer to the mean than is the predictor. If the student's ability score is not exactly average, the achievement score can be expected to fall between the ability score and the mean.

Since discrepancy formulas are concerned only with achievement lower than ability, there are two situations to consider. If the student's ability measure is higher than the mean (for instance, 100 for most intelligence tests), the expected achievement level without learning disabilities or other damaging influences will fall between the mean and the ability score, or above 100 and lower than the ability score. If the ability score is lower than the mean, the normally expected achievement score will again fall between the ability score and the mean, higher than the ability score. In this instance, the discrepancy would be from an expected achievement score higher than the ability score. The lower the ability score, the more the expected achievement score will exceed it. The amount of regression toward the mean is determined by the correlation between the two measures. The lower the correlation, the greater the regression. Contrary to general belief, the correlation between measures of ability and achievement is modest. In one example, Hammill and McNutt (1981) reviewed all correlational studies between reading and other variables in 25 journals from 1950 through 1978. Their meta-analysis found, among other things, that the median correlation between WISC and WISC-R Full Scale IQs and read-
Continued from page 7

Yes, Virginia, there is a Severe Discrepancy Clause, but is it Too Much Ado About Something?

5. It is highly improbable that the most valid measure of ability for a learning disabled student is the Full Scale IQ (GCA, GIA, etc.). The same learning disabilities that depress achievement are likely to depress scores on at least some parts of ability tests as well (Mark Penalty). The team is evaluating discrepancies between achievement and levels of intellectual ability, not overall intelligence. Consequently, the team should often use measures of intellectual ability other than full scale scores (IQs, GCAs, GIAs, etc.).

6. The achievement test scores may not be the most valid measures of a student's achievement. Some students perform poorly on standardized tests, but succeed in class and do not require special education services. Other students may do well on individual, standardized tests, but still have legitimate special education needs because they cannot function in class without a program of special education. It is not reasonable to use the standardized test scores as the only measures of achievement. Measures of classroom performance must also be considered.

7. Tests are not interchangeable. Two intelligence tests, even from a District's "approved" list (another unwise practice, unless exceptions are allowed for special circumstances), are not likely to yield the same score for the same student. In fact, the same test probably will not give the same score to the same student twice in succession. Consider, for example, a hypothetical student with a significant strength or significant weakness in fluid reasoning (Gf). The strength or weakness would have almost no direct effect on the Wechsler Scales, but would be central to one-seventh of the WJ-R or WJ III, and one-third of the DAS (Flanagan, McGrew, & Ortiz, 2001; Flanagan & Ortiz, 2001; McGrew & Flanagan, 1998). Differences among content, formats, and scores among academic achievement tests are even more extreme. Discrepancy formulas can lead to bizarre efforts to find combinations of IQ and achievement tests that will confirm someone's belief that the student does, or does not, have an educational disability.

Clearly, the mathematical process of rigidly determining severe discrepancies is not a simple matter. In fact, we do not believe that it is worth pursuing. Rigid adherence to any formula will often violate the most important consideration of all: to be considered to have an educational disability, a student must require a program of special education. In this respect, learning disabilities are no different from other educational disabilities. For instance, one student with an intellectual disability (mental retardation) with a certain IQ score will require special education services and will have an educational disability. Another student with the same score on the same test will not. The same consideration applies to students with various Snellen vision ratings or better ear average hearing ratings. Some students with a given score will require special education and some will not. Only those who require a program of special education have educational disabilities. No mathematical formula should be used to deny services to students who truly require a program of special education. No mathematical formula should be used to falsely "label" a student who does not require special education services, even if the student may require accommodations under PL 94-142 (U.S. Congress, 1975). "The determination of an educational handicap is a Team decision, not an exercise in arithmetic" (Willis & Dumont, 1998, p. 112).

The question as to whether a child does or does not have a severe discrepancy is not as straightforward as it might appear. While mathematical
There are other issues that may be argued before a hearing officer when a child does not appear to meet state criteria. Brody v. Dare County, a North Carolina case argued by Peter Wright, an eminent attorney in the area of special education law, included the following (Findings of Fact provided by a state hearing officer):

Dyslexia is a learning disability for the purpose of the Individuals with Disabilities Education Act, 20 U.S.C. Section 1400 et seq. (IDEA), and a child with special needs pursuant to North Carolina’s Special Education Act, G. S 11C, Article 9 (State Act).

The “Matthew Effect” refers to individuals, who have difficulty learning to read, and whose reading problems are not remediated effectively early in their school career. Because of this these children often remain significantly behind peers in reading skills. If an individual fails to learn well in first, second, and third grade—which are critical grades—then these individuals do not read the amount of material that is necessary to continue to develop good reading skills.

• Q. When determining a severe discrepancy, doesn't the child have to have at least average ability?

There are some misconceptions that exist about the regulations regarding learning disabilities and a level of cognitive functioning, some that are even codified into state regulations. IQ is not an exception to eligibility. Whether a child’s IQ is high or low, the team must still consider whether he or she meets the eligibility criteria in 300.541. (Letter to Ulissi, OSEP 1992). There is no place in IDEA that necessitates an IQ level for classification as a child with a learning disability. IDEA ’97 Final Regulations, Subpart A §300.7 (c)(10)(ii) does note:

(ii) Disorders not included. The term does not include learning problems that are primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

On first glance, this section would appear to exclude, among others, any child who is cognitively impaired (mentally retarded). However, the clause contains the operative phrase "primarily the result of . . . " A child with low cognitive scores may in fact be a child with a disorder in a basic psychological process and more appropriately identified as LD than MR. Also, a child with mental retardation may additionally have a disorder in a basic psychological process that depresses the child’s achievement in one academic domain significantly below even the low expectations from the child’s measured levels of intellectual ability and below the child’s levels of achievement in other domains. Such a child would have both mental retardation and a specific learning disability, because the excessively low achievement in the one domain was caused by the processing disorder: that particular “learning problem” was not "primarily the result of . . . mental retardation."

States have run afoul of OSERS when they have written state regulations that are more restrictive than the federal criteria in CFR Section 300.541. Wisconsin’s original definition of a specific learning disability was confusing, and led some teams to conclude that otherwise qualified children weren’t eligible because they had IQs below 90, and other teams to conclude that otherwise eligible child under the federal statute was not eligible in Wisconsin because there was only one area of discrepancy.

An audit of their public school systems showed that while most schools were in compliance, some teams believed (1) children with IQs...
below 90 could not qualify for services, or (2) that if children only qualified in one area they could not be served, or (3) that if a child had only a discrepancy, s/he was automatically eligible without first determining a need for special educational services. They issued the following statement to the local educational agencies (LEAs) as part of their corrective action plan (a plan necessary for them to continue receiving special education dollars from the federal government, ED having frozen their funds):

“If an M-team determines that a child has a significant discrepancy between functional achievement and expected achievement in one or more of the areas listed at 34 CFR 300.541 and needs special education and related services because of that significant discrepancy, the child may not be excluded from LD eligibility because: 1) The child's intellectual capacity is below a particular level (unless the child is determined to have a cognitive disability), or 2) the child has a significant discrepancy in only one of the academic or readiness areas identified at 34 CFR 300.541.”

Wisconsin’s interpretation, unchallenged by OSERS, is summarized in the following paragraph from their letter to Thomas Hehir, then OSERS’ director:

“M-teams in Wisconsin generally recognize that the criteria contained in both the federal regulations and the state rules must guide an evaluation, but do not direct an M-team to make a finding of LD eligibility or ineligibility for a particular child. Both the state and federal criteria are permissive in that they require evaluation teams to consider certain eligibility criteria, but they do not require an evaluation team to reach a conclusion solely because the child meets or fails to meet those stated criteria. The rules require evaluation teams to consider the criteria and

the performance of the child against those criteria, but they also require the evaluation teams to use professional judgment in making individual eligibility determinations.”

Q. So what are the issues that a MDT might consider when deciding whether a child qualifies as a child with a specific learning disability?

It is probably not possible to list ALL the factors that a team might consider in reaching a decision, but court decisions and OSERS letters have given us some suggestions.

1. Neither a low IQ score nor a high IQ score (see above) may be used to exclude a child from consideration as LD. http://www.ldanatl.org/bulletins/AC_1_96.html

2. The absence of a numerical severe discrepancy cannot, alone, be used as the criterion for excluding or for including a child from or in special education.

3. It is "generally" appropriate for the multidisciplinary team to include in its written report (to determine eligibility) information regarding "outside or extra" instructional help or support which "may indicate the child's current educational achievements reflects the service augmentation, not what the child's achievement would be without such help." Within context, for example, a child need not have failing grades, if he or she is passing only as a result of special service or support such as tutoring twice a week or a parent who spends three to five hours with the child on homework each evening.

4. The team should understand and take into consideration both the Mathew Effect (Stanovich, 1994) and the Mark Penalty. If there is prior evidence of higher IQ, and present testing shows a decline (Mathew Effect) that results in the child being ineligible, the team may consider whether the disability may have resulted in significantly different learning experiences which have negatively impacted the scores. IQ scores, depressed by a disorder in basic psychological processes, cannot be used in any ability-achievement comparisons. If the team believes that the same disorder in basic psychological processes that has caused low achievement has also depressed an IQ score (Mark Penalty), the team may consider this in the explanation of no discrepancy.

5. If the multidisciplinary team determines that the assessment measures do not accurately reflect a discrepancy between achievement and ability, the team should state in writing the assessment procedures used, the assessment results, the criteria applied to judge the importance of any difference between expected and current achievement, and whether a substantial discrepancy is present that is not correctable without the provision of special education.

Additionally, here are a few things that a team can consider in determining if there is a severe discrepancy:

A. How much help did the student receive (within the context of OSEP’s letter, referring to the help provided by parents, but certainly a standard that could be extrapolated to children who, for example, no longer had a severe discrepancy after three years of intensive service)?

B. Did the child’s IQ scores go down as a result of the Matthew Effect or Mark Penalty? [In theory, if the team believes that the child’s FSIQ went down as a result of his or her having failed to thrive academically, that is, the child wasn’t exposed to the same things as other children of the same age and intelligence—then it can infer higher potential, especially if previous test scores were higher. There is a potentially interesting Catch 22 (Heller, 1961) there; the team also would have to rule out inappropriate instruction as being primarily causal.]
C. Is there another way of documenting the federal standard other than the state’s formula? [If so, the team could consider the child for eligibility. While we do not have documented instances of this, failure to consider other standards could leave one vulnerable under Section 504.]

Other jurisdictions outside the 9th Circuit Court have not accepted the limitations imposed by Larry P v. Riles. However, when teams doubt the validity of an IQ score based on one or more of the reasons noted above, we advocate that the team consider the same factors that California would review in making an eligibility determination (i.e., using data from the child’s educational and social history as well as adaptive behavior data to make inferences about the child’s potential). The total evaluation and determination of eligibility for Special Education must be an integrated process. Again, SLD identification is a professional judgment by a team, not an exercise in arithmetic.

According to IDEA §300.540:

The determination of whether a child suspected of having a specific learning disability is a child with a disability as defined in §300.7, must be made by …the child’s parents and a team of qualified professionals …

8. The federal regulations require that the child demonstrate a disability in a basic psychological process which would include (but is not restricted to) a perceptual disability. This is basic to the identification process, but it is not as complicated as some would make it. While the idea that there must be a processing disorder that is inherent to or intrinsic to the child seems complex, in fact for most children establishing such a disorder can be as simple as identifying the child’s learning problem and narrowing in on other, highly probable causes.

"[1] It is not necessary for the multi-disciplinary team to demonstrate or measure the existence of a basic disorder in psychological processing in order to determine that a child has a specific learning disability. Rather, if a psychological processing disorder exists, it could manifest itself through a variety of symptoms that could be observed such as hyperactivity, attention problems, concept association problems, etc. See 34 CFR [Section] 300.542. The end result of these symptoms is a severe discrepancy between achievement and ability.” [Letter to Kennedy, IDELR 16 EHLR 1082, (OSEP, 1990).]

For example, if the child has trouble with word identification, it would be reasonable to infer that, if the child’s learning problems are not a result of cultural, environmental, or economic factors or one of the other exclusionary factors listed above, that the child might have a deficit in phonological processing, fluent retrieval, and/or oral vocabulary—all basic psychological processes within the meaning of the Regulations. These hypotheses could be tested quickly and easily by observation or by specialized tests. So, Virginia, to summarize our points, we offer:

**LOGICAL STEPS IN DETERMINATION OF A SPECIFIC LEARNING DISABILITY**

1. **Is there a problem with academic performance?** Problems may be subtle or difficult to document, but if there are no academic problems at all, there is no educational disability. A problem with an important life function other than academic performance might trigger identification under Section 504 of P.L. 93-112 or the Americans with Disabilities Act (ADA). Pay close attention to reports of problems that do not result in low grades even though they interfere with learning. For example, the teacher might already be providing an informal program of special education; grades might be based 25% on attendance, 50% on simply turning in homework regardless of quality; and 25% on class participation; or grades might be based on an erroneous perception of the student’s academic potential.

A. Does the student have low scores on group or individual achievement tests?
1. Look at any history of test scores. Be cautious, though, with tests that are used so frequently that the expected growth from test to retest is less than the 90% confidence band or even the SEM. Check the tables.
2. Look at the pattern of strengths and weaknesses on the test scores. Some group tests offer item analyses. Even though the norm-referenced tests do not function well as criterion-referenced measures, those analyses may contain useful information.

B. Is the student receiving low or failing grades in a class?
1. Again, track the history of class grades.
2. Try to determine the basis for the student’s grades. High grades might be based on special marking considerations.

C. Is the student working much too hard or much too long to earn adequate grades?
1. Parents may be the best source of this information. A parent interview is essential. We need to know also what the parents would like to learn from the evaluation.
2. Be sure to interview the student. Sometimes it helps to obtain a copy of the report card and discuss it in detail with the student. What does the student want to learn from the evaluation?

D. Is the teacher making extraordinary adaptations or accommodations for the student?
1. Teacher interviews are essential. We need to know what has been
done, what is being done, how well those interventions have worked and are working, and what specific things the teachers would like to learn from the evaluation.

2. The classroom observation is often more useful for observing the teaching and the environment than for observing the student.

E. Is there a notably deficient specific area of performance (e.g., tests, homework, note-taking, etc.)?

F. Is there another indication of insufficient academic performance?

2. Are there one or more disorders in basic psychological processes involved in understanding or in using language, spoken or written? [See http://alpha.fdu.edu/~dumont/basic_disorders.htm] This step follows next in a logical sequence, but determination of any disorder(s) may not be clear until completion of psychological, educational, speech and language, occupational therapy, physical therapy, vision, hearing or other evaluations. There should be multiple, convergent confirmations of any disorders.

A. Can each disorder be observed or inferred from academic performance?

1. Again, consider all aspects and all measures of academic performance.

2. Looking for possible cause-and-effect relationships between basic processes and academic performance. There needs to be a real-life connection between the hypotheses and what is actually happening with the student's performance in school.

B. Can each disorder be documented through assessment?

1. Once we have documented the deficient achievement and are looking for possible reasons, it becomes more permissible to use poorly normed and completely informal measures and observations. Formal assessment of ability and achievement levels needs to be done, at least in part, with extremely well-normed, reliable instruments that are valid for their intended purposes, but exploring within the area of deficient achievement may (and sometimes, given the state of the art, must) be done with less statistical rigor. The disorders need to be demonstrated clearly, reliably, and convincingly, but not always as test scores. The severity of a learning disability is measured by the severity of its impact on achievement, not by the severity of any basic-process disorder.

2. The McGrew, Flanagan, and Ortiz integrated Cattell-Horn-Carroll (CHC) Cross-Battery Approach is a very useful framework for considering many, though not all, basic-process disorders [See http://home.att.net/~gfgc/index.htm].

3. Can the team make a logical argument that each identified disorder manifests itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations? It is not enough simply to specify deficient achievement and a disorder. There needs to be a logical, cause-and-effect relationship between the two.

A. As noted above, we need to demonstrate how the purported basic-process disorder is impairing the carefully documented achievement area. This demonstration will require a thorough analysis of the student's academic skills. A low test score or low class grade is not enough. We need to show the mechanisms operating in the deficient achievement area(s). Examples of misaligned math problems worked left-to-right and bottom-to-top might, for instance, demonstrate the effect of a visual perception problem on math. The assumption that a visual perception problem impaired listening comprehension might be more difficult to justify unless, for example, we could show how deficient visual imagery was interfering with the listening comprehension.

B. Research evidence can be cited to show relationships between certain basic processes (e.g., phonological abilities or rapid naming) and certain areas of achievement (e.g., reading decoding). [See http://home.att.net/~gfgc/recent.htm for some examples.]

C. Some clearly identifiable disorders have no discernable effect on achievement. Simply finding a disorder does not establish a learning disability (e.g., one author's (JOW) severe rhythm disorder impairs his singing, dancing, and clapping in time to music, but the effect on academic achievement is trivial, only diminishing his appreciation of poetry).

D. It is the disorder in the basic psychological process that distinguishes a specific learning disability from the disabilities and disadvantages ruled out in the federal regulations [(300.7(c)(10)] for learning disabilities (“...learning problems that are primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.”)

E. It is essential, as much as possible, to distinguish learning disabilities from dyspedagogia and apedagogia [300.541(1) "The child does not achieve commensurate with his or her age and ability levels in one or more of the areas listed in paragraph (a)(2) of this section, if provided with learning experiences appropriate for the child's age and ability levels" (emphasis ours).]

4. What is the best estimate of
the student’s actual intellectual ability? See Mark 4:25. The team must not allow a psychological processing disorder to depress estimates of both intelligence and achievement and then mindlessly conclude there is no discrepancy between the two. For example, verbal and visual/spatial learning disabilities, respectively, will depress verbal (Gv) and visual, spatial (Gv) intelligence measures. For another example, a disorder in quantitative knowledge (Gq) would depress the WISC Arithmetic and Verbal IQ scores and DAS Sequential & Quantitative and Nonverbal (fluid) Scale scores. Logically, the intelligence test should be chosen only after the basic process disorders have been delineated. The McGrew, Flanagan, and Ortiz integrated CHC Cross-Battery Approach can be a very useful framework for considering intellectual abilities [See http://home.att.net/~gfc/index.htm.]

A. Which scales, factors, or subtests on intelligence tests are likely to be depressed by the disorder or disorders?
B. Which intelligence test, scales, or factors would be likely to yield an estimate of actual intellectual ability uncontaminated by the disorder or disorders?
C. What is the best estimate of the student’s actual intellectual ability based on those measures?
D. Have we considered at least all of the broad abilities in the McGrew, Flanagan, and Ortiz integrated CHC theory? It is not prudent, for example, to use a test, such as the WISC-III, that omits fluid reasoning unless we supplement it with a measure of that ability.

5. Is there a severe discrepancy between the student’s level of intellectual ability (4. C.) and the student’s achievement in one or more of the following areas? Remember that achievement and ability may be assessed by means other than test scores (1. B. – I. F.). Achievement tests must be chosen thoughtfully. For example, a brief achievement test is not a valid measure of academic performance for a student with a short attention span, and an untimed, silent reading test will not pick up problems with reading fluency. Do not obsess over formulas. Some data will not fit formulae. The team must apply reasoned, professional judgment, not simply indulge in an exercise in arithmetic. By our interpretation of federal law and by most state laws, it is not lawful to deny services to a student who truly has a learning disability simply because of the results of a statistical exercise. [See http://alpha.fdu.edu/~dumont/riley_v_ambach.htm.]. A statistical comparison of ability and achievement must use only one set of norms (e.g., national age or grade) [See http://alpha.fdu.edu/~dumont/age_vs_grade_based_scores.htm.] and should use predicted achievement scores rather than simple differences [http://alpha.fdu.edu/~dumont/Severe_Discrepancy_Discrepancies.htm#SD2, http://alpha.fdu.edu/~dumont/Determining_predicted_ach.htm, and http://home.att.net/~gfc/psych0101b.htm].

Remember that these achievement areas have many components, including, for example, vocabulary or factual knowledge, fluency, independence. Few, if any, achievement tests cover all aspects of the requisite skills. Do not use tests on which the student receives very low or nearly perfect raw scores, but find tests on which the student passes and fails several items [http://alpha.fdu.edu/~dumont/McGee.htm.]

6. Are the discrepancies caused primarily by the disorders? There is absolutely nothing in IDEA to suggest that a student cannot have a learning disability in addition to other disorders. However, the particular discrepancy (“learning problems”) in question must not be primarily the result of a vision, hearing, or motor disability, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage [300.7 (c) (10) (ii)], even if one or more of those disorders or disadvantages may be causing other, separately identified learning problems. For example, a child with mental retardation might also have a specific learning disability in math with extremely low achievement severely discrepant from low predicted achievement because of a disorder in working memory. Similarly, a deaf or blind child might have unusual difficulty learning American Sign Language or Braille because of spatial perceptual weaknesses. If we have been careful in our identification and analysis of the disorder(s), we should be able to separate them and their effects from the effects of disadvantages and other disabilities.

7. Does the student require special modifications of, or accommodations in, the educational program in order to achieve at levels commensurate with age and ability (4. C.)? Here is the crucial issue for identification under Section 504 or the ADA. The needed accommodations or modifications should be more than we would routinely ask of a teacher of moderate skill, experience, and dedication.

8. Does the student require a uniquely designed program of special instruction in order to achieve at levels commensurate with age and ability (4. C.)? This is the crucial issue for identification of an educational disability. If the student does not require a uniquely designed program of special instruction, but meets the other criteria, the identification should probably be under Section 504 rather than the Individuals with Disabilities Education Act.

References

Continued on page 15
School psychology is still searching for an identity embraced by all. School psychology’s identity appeared to have reached an impasse almost 20 years ago, resulting in a dilemma that was written about and discussed many times since. It is clear from Jack Bardon’s writings (Bardon, 1982a, 1982b, 1983) and the accompanying comments and responses that the ideological and philosophical differences between the American Psychological Association (APA) and the National Association of School Psychologists (NASP) we see today were just as vast between the two organizations of APA’s psychology “specialty in search of an identity” (Bardon, 1983) and on the other hand we find NASP’s view of school psychology as it’s own profession with the entry level for independent practice being at the non-doctoral level (Bardon, 1983) and on the other hand, there is APA’s psychology “specialty in search of an identity” (Bardon, 1983) and on the other hand, we find NASP’s view of school psychology as being it’s own profession with the entry level for independent practice being at the non-doctoral level (Bardon, 1982b). APA’s School Psychology Division (Division 16) weighed in heavily in the mid-1990’s in this discussion (e.g., see president’s columns from 1995 to 2000 in The School Psychologist). Since the discussions stirred by Bardon and others began, there were advances in consolidated meetings and agreements between the two organizations of APA and NASP with these underlying differences that seemed to have been set aside, but were never able to ignore.

Division 16 (D16) surveyed its membership in 1995-1996 to determine its view on defining our profession. Based upon those results and additional discussions, the Executive Committee for D16 forwarded a motion that detailed the role of D16, defined it’s mission and stated that the “...most appropriate name for the practice specialty is doctoral school psychology” (Illback, 1996, p. 41). The APA/NASP Interorganizational Committee (APA/NASP IOC) has struggled with these philosophical issues while at the same time working to strengthen school psychology as both a profession and as a specialty of the profession of psychology. In the early 1980s, there was even a joint accreditation project between both organizations (APA/NASP to Try Joint Accreditation, 1982). Similarly, there was mutual work and agreement between APA and NASP due in large part to D16 and the work of the APA/NASP IOC on the development of APA’s Specialty Guidelines for School Psychology (1998). Recently, Rick Short (2000) discussed the two cultures, similar to Bardon’s (1982a) discussion and brought the issues back to the forefront as NASP completed the process of renewing their guidelines for practice and training. Unfortunately we find that although there are things that the two organizations can mutually agree upon, the basic philosophies do impede complete agreement and successful collaboration. The recent adoption by NASP of their Standards for the Credentialing of School Psychologists, Standards for the Training and Field Placement Programs in School Psychology, and Standards for the Provision of School Psychological Services (NASP, 2000) have resulted in accentuating the differences between NASP and APA. No longer do we see minor differences with similar overall goals. NASP has clearly diverged from the previously agreed upon mutual accepted standards. The new NASP documents are clearly different in training, accreditation, credentialing/licensing, and practice standards from APAs.

One result of NASP’s move away from the status quo has been to bring to flame the smoldering aspects of this identity issue in school psychology. Is school psychology a "Separate Profession Culture" as defined by Short (2000) and unique among helping professionals or are we a specialty of professional psychology with training and experiences in schools designed to extend beyond the walls of school buildings? The search for an identity continues.

References
and for discussion of cases. Our first debate took place within a cognitive assessment seminar offered in the fall of 1997. Students were asked to present positions for or against the use of cognitive assessment. Each student was assigned an aspect of the debate to cover. The subtopics included item bias, differential predictive validity, construct validity, etc. First, a student entered an initial position statement, either pro or con, on the assigned subtopic. Then two days later, after a critical analysis of the position statement of the student’s opponent, a rebuttal was posted. This mirrored the rebuttal phase of a conventional face-to-face debate, except that the students could read and analyze the sources used by the opponent in the initial position statement. Finally, students in the class reviewed all the statements and rebuttals, and developed their own personal statements, which were shared in the public forum that the class had access to.

Another use of web forums is to collaborate on cases. Indiana University interns from across the country, some in urban settings, some in rural settings, submitted cases to an electronic forum. Teams of first year school psychology graduate students reviewed the cases, went to the professional literature, proposed interventions supported by the empirical literature and were critiqued by peers and faculty. Their presentations were posted in the electronic forum so that interns had the opportunity to react.

I am not advocating that electronic communities replace conventional place-based professional associations. However, members of what have been place-based associations may now interact in new ways. New electronic communities will supplement conventional professional associations in ways we cannot even imagine. In other words, we are no longer place-bound or time-constrained when interacting as professionals.

I look forward to 2001 as a time when we will explore the possible uses of electronic media to develop professional communities where colleagues can interact, share resources, and debate the merits of ideas. We have much to gain from the experiences of colleagues across the world. The new Internet-based technologies will permit us to form alliances and professional communities that hitherto would have been unworkable. The old adage, "Think globally, Act locally," might be turned around to form a new challenge, "Think globally, Act locally."

Continued from page 2
President’s Message: Professional Communities of the Future

Continued from page 13
Yes, Virginia, there is a Severe Discrepancy Clause, but is it Too Much Ado About Something?


Larry P. v. Biles, 793 F.2d 969, 974 (9th Cir. 1984).


Comments, thoughts, and opinions regarding this article for the Comment Section of The School Psychologist should be e-mailed to: Reddy@alpha.fdu.edu.
I am often asked what I think was the most important event in the history of school psychology. Since school psychology's history is little more than 100 years old, and since this is officially the beginning of the 21st century, now is as good a time as any to register my opinion.

There are many worthy nominations. Among them would have to be the founding of the first psychological clinical in 1896 or the Department of Scientific Pedagogy and Child Study in the Chicago Public Schools in 1899; the enactment of compulsory schooling laws; the development of special education; the development of the age-based Stanford Revision of the Binet Simon Scales in 1916, or the deviation IQ concept of the Wechsler scales a few decades later; Gesell’s practice as a “school” psychologist 1915-1919; the widespread acceptance of normative psychoeducational assessment in school settings; the development of a school psychology training program at New York University in the 1920s; the development of state education agency credentials for practitioners in the 1930s; the founding of the Ohio School Psychologists Association in 1943; the reorganization of the American Psychological Association (APA) in 1945 with a separate division for school psychologists; APA's first code of ethics in 1953; the Thayer Conference in 1954, or the Spring Hill and Olympia Conferences in 1980 and 1981, respectively; the publication of Psychological Service for School Problems in 1930, or the Journal of School Psychology in 1963 and Psychology in the Schools in 1964; the founding of the National Association of School Psychologists (NASP) in 1969; APA accreditation in school psychology and the NASP-NCATE relationship since the early 1970s; the development of standards documents by APA and by NASP; the approval of learning disability as a special education category in the late 1960s; PL 94-142 (Education of All Handicapped Children Act, 1975) and PL 93-380 (Family Educational Rights and Privacy Act, 1974); the widespread acceptance of the importance of mental health and of therapy in the past 40 years; NASP's program of National Certification in School Psychology in 1988; the emphasis on consultation-based alternative delivery systems since the late 1970s and nontraditional assessment emphases of the 1990s; and finally, official specialty recognition granted to school psychology through APA in 1997.

Of course, all of these have made important contributions to the development of school psychology in the 20th century and no single event could account for the current condition of school psychology. Nevertheless, I have forced myself to choose the most important. After considerable deliberation I decided that the following was the most important single event: The fact that early psychologists chose to become employees of the schools rather than practitioners to the schools. The choice has had bittersweet outcomes. The increasingly frequent decision of early practitioners to accept employment with school districts led quickly to the following circumstances in many aspects of practice. Some have been adverse and they are presented here in no particular order:

1. The school psychologist became an employee of the school board rather than a professional whose client was the school board.
2. The purpose and focus of the school psychologist's work could be primarily determined by the employer.
3. Resources available for practice could be limited to the resources of the employing school district.
4. Scope of practice could be limited to school-related learning problems, undervaluing related mental health and family concerns.
5. Scope of practice could be limited to school district personnel and physical facilities, undervaluing parental and community involvement.
6. Supervision often would be provided by persons having no prior training or experience in psychology.
7. The school psychologist would practice in a setting where non-doctoral credentials would be desired and acceptable.
8. The political status of the school district and public education could influence the work of the school psychologist.
9. The school psychologist became most closely connected to special education.
10. The school psychologist might need to be a teacher before being a psychologist.
11. Practitioners would often be perceived as pupil personnel services workers rather than professional psychologists or health service providers.
12. Regulation of training and credentialing would become the responsibility of education-related groups and authorities, in addition to or exclusive of, those in psychology.
13. Like the separation of the schools from the broader community, school psychology would often be separated from the broader field of professional psychology.

The widespread employment opportunities in the schools looked like a Garden of Eden to early practitioners. Perhaps picking the fruits of school district employment became the field's original sin. I have deliberately emphasized the controversial aspects of this event. For much of the 20th century, school psychologists complained about, atoned for, and tried to alter these conditions, many of which helped to shape long-standing conflicts such as the doctoral and non-doctoral issue, or the...
gatekeeper role in special education. Many of the other nominated historical events are themselves outgrowths of the fact of school employment (e.g., NASP-NCATE relationship, SDE credentialing). Of course, there have been numerous advantages to being school district employees: office space, secretarial and communication services, materials and equipment, professional status in the system, fringe benefits such as retirement and health insurance, relatively good salaries, steady employment, 10 month contracts, direct access to serving children, etc.

History is the interpretation of events. For what it is worth, this is my interpretation of the most important event in the history of school psychology: It shaped us into "school" psychologists instead of psychologists in the schools (a controversy of its own). I am interested in your interpretations. Please send them to me at Department of Psychology, 202 Psychology Building, UM, Memphis, TN 38152-3230, or <tom-fagan@mail.psyc.memphis.edu>

Internships

1 a copy of this article will also appear in the March 2001 Communique.
Call for Nominations:
Senior Scientist Award

The Division of School Psychology (Division 16) of the American Psychological Association requests nominations for the Senior Scientist Award. This award is presented to school psychologists who throughout their careers have demonstrated exceptional programs of scholarship that merit special recognition. This is not an award necessarily for the amount of writing done by a scholar, but rather for a sustained program of outstanding theoretical and research activity. Nominees must be: (a) either 20 years past the granting of their doctoral degree or at least 50 years old by December 31, 2001, and (b) a Fellow, Member, or Associate of Division 16. The award recipient will be asked to prepare an address for the Division to be presented at the APA annual convention, submit a manuscript based on that address to School Psychology Quarterly (the Division 16 journal), and serve on a committee to select subsequent award winners. Anyone, including a candidate him or herself, may nominate a school psychologist for the award. Five sets of materials should be submitted for each nominee, including a vita, 3-5 supporting letters, and five major papers or publications. Send all nomination materials by April 1, 2001 to the chair:

Dr. Jerome Sattler
Emeritus Professor, Department of Psychology
San Diego State University
5500 Campanile Drive
San Diego, CA 92182-4611
619-594-6231 (phone)
619-594-1332 (fax)
Jsattler@sunstroke.sdsu.edu

Call for Nominations:
Lightner Witmer Award

The Division of School Psychology (Division 16) of the American Psychological Association requests nominations for the Lightner Witmer Award. This award is presented to school psychologists who have demonstrated exceptional scholarship early in their careers. Continuing scholarship, rather than a thesis or dissertation alone, is the primary consideration in making the award. Nominees must be: (a) within seven years of receiving their educational specialist or doctoral degree as of September 1, 2001, and (b) a Fellow, Member, Associate, or Student Affiliate of Division 16. A person does not need to have a doctoral degree to be eligible. The award recipient will be asked to prepare an address for the Division to be presented at the subsequent APA annual convention, submit a manuscript based on that address to School Psychology Quarterly (the Division 16 journal), and serve on a committee to select subsequent award winners. Anyone, including a candidate him or herself, may nominate a school psychologist for the award. Five sets of materials should be submitted for each nominee, including a vita, 3-5 letters of support, reprints, and other evidence of scholarship. Send all nomination materials by April 1, 2001 to the chair:

Dr. Christopher Skinner
Psychoeducational Studies Unit
The University of Tennessee
438 Claxton Addition Building
Knoxville, TN 37996-3400
W (865) 974-8403
H (865)588-9501
cskinne1@utk.edu

Call for Nominations:
Jack Bardon Distinguished Service Award

The Division of School Psychology (Division 16) of the American Psychological Association requests nominations for the Jack Bardon Distinguished Service Award. This award is presented to mature school psychologists who throughout their careers have demonstrated exceptional programs of service that merit special recognition. This award is given for accomplishments relating to: (a) major leadership in the administration of psychological services in the schools, (b) major contributions in the formulation and implementation of policy leading to psychologically and socially sound training and practice in school psychology, (c) sustained direction and/or participation in research that has contributed to more effective practice in school psychology, and/or (d) the inauguration or development of training programs for new school psychologists or for the systematic development of in-service training for psychologists engaged in the practice of school psychology. The award recipient will be asked to prepare an address for the Division to be presented at the subsequent APA annual convention, submit a manuscript based on that address to School Psychology Quarterly (the Division 16 journal), and serve on a committee to select subsequent award winners. Anyone, including a candidate him or herself, may nominate a school psychologist for the award. Five sets of materials should be submitted for each nominee, including a vita, 3-5 supporting letters, and five major papers or publications. Send all nomination materials by April 1, 2001 to the chair:

Dr. Jerome Sattler
Emeritus Professor, Department of Psychology
San Diego State University
5500 Campanile Drive
San Diego, CA 92182-4611
619-594-6231 (phone)
619-594-1332 (fax)
Jsattler@sunstroke.sdsu.edu

Continued on page 19
address for the Division to be presented at the subsequent APA annual convention, submit a manuscript based on that address to School Psychology Quarterly (the Division 16 journal), and serve on a committee to select subsequent award winners. Anyone, including a candidate him or herself, may nominate a school psychologist for the award. Two sets of materials should be submitted for each nominee, including a vita, supporting letters (minimum of three), and other appropriate supporting documentation. Send all nomination materials by April 1, 2001 to the chair:

Ronda C. Talley, PhD, MPH
Executive Director and Professor
Rosalynn Carter Institute for Human Development
Georgia Southwestern State University
800 Wheatley Street
Americus, GA 31709
Talleyrc@aol.com
229-928-1234 (phone)
229-931-2663 (fax)

Call for Nominations:
Outstanding Dissertation in School Psychology Award

The Division of School Psychology (Division 16) of the American Psychological Association requests nominations for the Outstanding Dissertation in School Psychology Award. This award is presented to a school psychologist who has completed a doctoral dissertation which merits special recognition and which has the potential to contribute to the science and practice of school psychology. Nominees must: (a) have successfully defended the dissertation between January 1, 2000 and December 31, 2000 and (b) be a Member or Student Affiliate of Division 16 at the time of receipt of the award (August, 2001). The award recipient will be asked to serve on a committee to select subsequent award winners, give an award presentation based on the dissertation at the subsequent APA annual convention, and submit a manuscript to School Psychology Quarterly (the Division 16 journal). Anyone, including a candidate him or herself, may nominate a school psychologist for the award. Four copies of the nominee’s vita and letters of support from at least two members of the dissertation itself should be submitted for each candidate, along with a copy of the dissertation. Send all nomination materials by April 1, 2001 to the chair:

Dr. Kimberly P. Ray
Child Development Clinic
2500 N. State Street
Jackson, MS 39216-4505
Fax: 601-984-2975
Phone: 601-984-5239
kray@ped.unsm.edu

Call for Nominations for Division 16 Fellows of APA

The Division of School Psychology requests your nomination of individuals for Fellowship status in APA. Nominations to initial Fellow status are reviewed by the Division’s Fellows Committee, and forwarded to the APA Membership Committee, which has the responsibility of making recommendations to the APA Board of Directors. The APA Council of Representatives then elects individuals to Fellow status upon recommendation of the Board. Nominees must hold a doctoral degree, have been an APA member for at least one year, be engaged in the advancement of psychology, and have at least five years of professional experience after the doctorate. Election to Fellow status requires evidence of unusual and outstanding contributions or performance in the field of psychology. Fellow status requires that a person’s work have had national impact on the field of psychology beyond a local, state, or regional level. Three letters of endorsement from current APA Fellows will be required in support. Anyone, including a candidate her or himself, may nominate a school psychologist as a candidate. Upon receipt of a nomination, necessary information will be sent to the candidate who will prepare and return a formal application with supporting material to the Division Fellows Committee. Please send nominations as soon as possible to the chair:

Dr. Marian Fish
School of Education
Queens College, CUNY
Flushing, NY 11367
718-997-5230 (phone)
718-997-5248 (fax)
Mfish@gc.cuny.edu
Is There Black Culture Learning Style?

Mary Levinsohn

The purpose of this paper is to explore the empirical base of literature around learning styles and culture as related to academic achievement. The field of learning styles has been examined by several researchers including Dunn and her associates (Dunn, 1984; Dunn, Beaudry, & Klavas, 1989; Dunn & Dunn, 1978; Dunn, et al., 1990). A brief review of the literature on learning styles will be examined in order to give a context for the general purpose of the paper. However, the primary focus of this review is on the interaction of learning styles and culture as related to academic achievement. The School Psychology Review article by Frisby (1993) will be used as a guide for analyzing the germane literature. Conclusions and general themes around the Black culture learning style (BCLS) will be generated.

Learning Style

According to Dunn et al. (1989), "learning style is a biologically and developmentally imposed set of personal characteristics that make the same teaching method effective for some and ineffective for others" (p. 50). They go on to describe that learning style indicates the setting or context where a person learns best. Some of the examples given include whether a person learns best alone or in a group; auditorily, visually, or tactically; and with varying degrees of structure. Dunn (1984) describes learning styles as falling into five categories: environmental, emotional, physical, sociological, and psychological.

The topic of learning styles and modalities has received much attention. As Kavale and Forness (1987) indicate, the efficacy of learning style to increase achievement holds great intuitive appeal. The research put forth by Dunn and her graduate students shows potential promise for the use of learning styles (e.g., Dunn, et al., 1989). However, most of the research linking instruction to a child’s learning style is only found in unpublished dissertations. These dissertations mainly come from the university where Dunn is a faculty member. That research field would be strengthened if more empirically based research articles were published in peer reviewed scholarly journals. In addition, the meta-analysis by Kavale and Forness (1987) shows that in the overall empirical literature, support for the modality assessment and instruction is less than promising. Nonetheless, research and interest in the relationship between learning styles and academic achievement has received attention in the professional literature.

Black Cultural Learning Style (BCLS)

According to Frisby’s (1993) account, interest in the effect of culture on psychological processes (e.g., learning style, cognitive style) peaked in the 1970’s. The cultural effects on learning style was promoted as an explanation for the underachievement of African-American children in the classroom. In Frisby’s compilation of the BCLS literature, he gives a comprehensive description of the BCLS elements. The elements fall into four main categories: world view, cognition, learning, and communication/social. The literature he reviewed indicates that substantial differences exist in these areas when comparing African (Black) culture and European (White) culture.

Frisby’s Five Assumptions

Based on his review of the BCLS literature, Frisby (1993) conveys five assumptions that he recognized as themes in this literature. He goes on to debunk the five assumptions. Three of the five assumptions will frame the analysis of the empirical literature that will be reviewed next. Taken directly from Frisby’s work the five assumptions are as follows:

1) Within America, Black and white culture are fundamentally incompatible (p. 541).
2) Characteristics of Black culture "determine" the learning style of black children (p. 543).
3) Learning style assessment is reliable and valid (p. 545).
4) Black learners are incapable of manifesting any behavior that is different from behaviors dictated by their identified cultural learning style; therefore, teachers are morally obligated to adapt their instructional practices to Black children’s identified learning styles (p. 547).
5) There are "culture-specific" educational prescriptions that are uniquely effective in increasing academic achievement for black students relative to whites (p. 548).

Review of Extant Literature

A search was conducted to find the empirical literature that examined the effect of BCLS on academic achievement. These articles will be summarized in both methodological and conceptual terms.

Stated Purpose of Studies

Bell and McGraw-Burrell (1988) conducted a study that compared the performance of Black children on tasks when presented in a varied and an unvaried format. The varied format was hypothesized by Boykin’s theory to be congruent with BCLS, whereas the unvaried format was aligned with white cultural learning style. The researchers also wanted to determine whether the Black students’ scholastic achievement level related to their performance on the task, as a result of the format presentation (e.g., varied or unvaried).

Other studies used a correctional design to determine patterns in learning styles of Black students. Jacobs’s (1990) study compared learning styles of Black high, average, and low achiev-
ing students. Likewise, Dunn et al. (1990) examined and compared differences between and among four ethnic groups on learning style. Howard et al. (1993) compared the learning style of preschool Black children to determine if the problem-solving strategies used systematically differed. Finally, Rech and Stevens (1996) identified variables that may affect math achievement for Black fourth and eight graders. They examined gender, attitude toward math, economic status, self-concept and learning style.

**Definition of Learning Style**

In Curry’s 1990 critique of the learning styles literature, she states that one concern with the research in this area is the lack of a conceptually clear definition of learning style. To address this point, all articles were reviewed for their explicit statement of what learning style meant in their particular research. In the Bell and McGraw-Burrell (1988) research, learning style was not explicitly defined. Instead, the definition of learning style was drawn from the Boykin model using varied versus unvaried task presentation. Dunn and her colleagues (Dunn et al. 1990) defined learning style as being "composed of unique reactions to environment, emotionality, sociological preference, physiological traits, and cognitive-psychological inclinations" (p. 69). Jacobs (1990) defined learning style based on the Dunn conceptualization.

Howard, Watson, and Allen (1993) and Rech and Stevens (1996) both solely examined the effect of one aspect of learning style (i.e., field independence/dependence). Howard et al. specified that field independence means that the person is analytical, able to perceive the figure from the background, and is attentive to detail. Field dependence means that the person processes information more globally, perceives the whole context, and conforms to context cues. Rech and Stevens (1996) did not clearly state their definitions of field independence/dependence. In summary, the clarity of the definition of learning style differed across studies.

**Sample Characteristics**

The studies used diverse methods to determine their samples. However, most studies chose to include only Black subjects (Bell & McGraw-Burrell, 1988; Howard et al., 1993; Jacobs, 1990; Rech & Stevens, 1996). Dunn et al. (1990) was the only study to include other ethnic groups (e.g., Chinese-Americans, Greek-Americans, Mexican-Americans).

Frisby’s (1993) Assumption #4

"Black learners are incapable of manifesting any behavior that is different from behaviors dictated by their individual cultural learning style; therefore, teachers are morally obligated to adapt their instructional practices to Black children's identified learning style." (p. 547).

These articles were analyzed using Frisby’s (1993) fourth assumption related to whether Black learners can adapt to a different learning style. In general, the empirical articles did not substantiate that Blacks were incapable of adapting to a non-Black culture learning style. This is in line with Frisby’s review. Many of the reviewed articles produced an interaction between learning style and another independent variable (Bell & McGraw-Burrell, 1988; Howard et al., 1993; Jacobs, 1990; Rech & Stevens, 1996). The BCLS theory would predict that regardless of other independent variables (e.g., task, achievement, and age), the efficacy of learning style would be a constant. In other words, the congruence between learning style and instruction should be a stronger predictor of success for Black children. However, the empirical literature did not substantiate this effect. For example, Bell and McGraw-Burrell (1988) reported a three-way interaction between learning style, achievement, and task. They found that high achieving Black children did better on an "ordered recall task," when presented in an unvaried format. BCLS would predict that all Black children, regardless of task or achievement, would excel when the task is presented in a varied format.

Frisby’s (1993) Assumption #5

"There are 'culture-specific educati onal prescriptions that are uniquely effective in increasing academic achievement for black students relative to whites" (p. 548).

For many in the education field, this final assumption of Frisby’s is the most compelling. Essentially, does using educational or instructional...
methods that are congruent with cultural learning style have a positive effect on academic achievement for Black children over white children.

The research reviewed for this paper did not advance the knowledge base in this area. Four of the five studies used only Black samples (Bell & McGraw-Burrell, 1988; Howard et al., 1993; Jacobs, 1990; Rech & Stevens, 1996). Clearly, the comparison with achievement with white children is not possible if only a Black sample is employed. The fifth study (Dunn et al., 1990) used a multi-ethnic sample (e.g., Mexican-Americans, Greek-Americans, and Chinese-Americans), but did not include an outcome measure of academic achievement or the modality if delivered instruction.

Interestingly, Rech and Stevens (1996) drew the conclusion at the end of their study that Black children would benefit from instructional practices that fit their learning modality. However, nowhere in their research do they manipulate instructional practices based on learning style. Therefore, drawing this conclusion is faulty and not substantiated by their research. This type of speculation and conjecture should be warned against when examining the BCLS research and any effect on academic achievement.

**Limitations of the Studies**

One major limitation seen across four of the studies was the use of a solely Black population. This decreased the ability to draw comparisons with white/majority culture. Many samples did not allow for external validity due to small sample size, unrepresentational samples, and poor description of sample characteristics. The correlational designs employed by many of the studies decreased the ability to determine causality. The studies would have been strengthened had an experimental manipulation been used. As noted earlier, the use of the assessments of learning styles was a major limitation of the studies. The reliability and validity of these measures need to be substantiated and included in the research write-up of studies using such assessments.

**Summary**

The purpose of this paper was to examine the learning styles literature, specifically as it relates to Black culture learning style (BCLS). First, this paper described the literature on learning styles and critiqued that empirical base. Next, Frisby’s (1993) article that examined Black culture learning styles was reviewed. From that article, a summary of the BCLS characteristics was provided. The articles were reviewed methodologically and conceptually. The Frisby article provided the conceptual framework from which to analyze the conceptual components of the BCLS literature.

In summary, several ideas became clear during the course of this research and analysis of learning styles and BCLS. First, the learning style literature is not empirically substantiated. As Kavale and Forness (1987) state, learning style makes a great deal of intuitive sense. However, the empirical support for appropriately assessing learning styles and the efficacy of using the learning style information to create better instructional environments has not held up in the scientific literature. Second, the empirical support of using instruction matched to a “Black cultural learning style” to increase academic outcomes for minority children is unsubstantiated. The research in the area of BCLS and increased academic outcomes is limited by its correlational nature, lack of experimental manipulation, unrepresentative samples, faulty learning style assessment measures, and lack of any substantiation that gearing instruction to a BCLS will change outcomes.

**References**


SASP News
The Newsletter of the Student Affiliates in School Psychology

SASP News Update
Gena N. Ehrhardt
Indiana State University
SASP News Editor

When we began this school year, SASP News had specific visions and goals. Two of the goals for the Communications Committee were to improve the communication among chapters, and to begin featuring writing and research of graduate students. In order to achieve these goals, students were invited to join the SASP News writing staff. I am pleased to announce that Melinda Stanley of Indiana State University will be serving as Layout Editor, and Caroline McKnight of the University of South Carolina will be a featured columnist writing about issues regarding school psychologists in hospital settings.

A new addition to SASP News is the featuring of topics important to the field of school psychology. In this winter edition of SASP News, the topic of intelligence testing is addressed. Additionally, SASP News is spotlighting various chapters and their local involvement. The chapters featured in this edition are Fordham University, University of Missouri-Columbia, and Georgia State University. Membership, listserv, and convention information are also included. Beginning January 2001, graduate students will be able to view web-based editions of SASP News at www.saspweb.org. As the Communications Committee continues to grow, I am optimistic about having new additions for each edition.

Returning to the featured topic, intelligence testing was selected for this edition for several reasons. A consistent component found in every school psychology training program is a course in intelligence testing. This course generally consists of the historical use of intelligence testing, the controversies in assessing intelligence, various theoretical perspectives which attempt to define intelligence, and finally, the administering and scoring of various intelligence tests. To practitioners in the field, intelligence testing demands most of their professional time. To students in training programs, intelligence testing transcends among several courses, including practicum and internship experiences. Therefore, presenting a forum for students to broach this controversial topic became a primary concern for the newsletter staff.

Due to its level of importance in the field, SASP News has included articles in the winter edition which address the subject of intelligence testing. The featured editorial piece composed by David Shriberg, president-elect of SASP, raises concerns regarding the use of intelligence testing for diagnostic labeling and whether such labeling is necessary for the provision of educational services. He challenges graduate students to determine their positions regarding diagnostic labeling in order to raise the consciousness of what we, as future school psychologists, will accept in professional practice. As an informative piece, columnist, Caroline McKnight, provides a suggestive list of accommodations and considerations when conducting assessments in a hospital setting. Caroline’s article informs graduate students about the unique testing considerations when assessing pediatric patients.

These articles are intended to promote discussion among graduate students and faculty. While the opinions regarding intelligence testing may vary from program to program, the topic, itself, is a commonality that is shared by us all. As developing professionals, it is important that we determine our position, and I encourage graduate students to form their own views on the topic and determine the role that intelligence testing will play in their professional practice. The amount of power that is associated with intelligence testing can be enormous. Therefore, it is important as future school psychologists to prepare for this awesome responsibility in a judicious manner in order to make conscientious decisions as to how such power will be used to best serve clients.

Check out SASP’s web page at: http://sac.uky.edu/~mdturn0/sasp.htm

The School Psychologist as Social Enforcer: Is This What We Went to School For?
David Shriberg
Northeastern University
SASP President-Elect

When I first entered graduate school in school psychology in the Fall of 1997, there were probably few students in the country more excited than I to be entering the field. Having been told all through my undergraduate days that clinical psychology was the only option for a “legitimate” psychologist, I came to Northeastern full of enthusiasm both to fight for greater recognition of school psychology as a field and also to find my own niche within it.

Now, as my graduate days (hopefully!) are beginning to come to a close and I am starting to experience the dreaded “real world” (I began working half-time as a school psychologist for two public elementary schools a few months ago), I am realizing that in my enthusiasm to educate the world about who a school psychologist is and what he or she does, I have forgotten one vital component of the job description. Nowadays, when somebody asks me what I do, I tell them that I have entered politics. And nowhere is the political function of my role as school psychologist more evident than in learning disability assessments, where my legal function is to administer an IQ test. I purposely use the phrasing “legal function” to reflect a change in the laws here in Massachusetts that has had a profound impact on the practice of school psychology. Some brief background information—starting in September, 2000, Massachusetts moved from a “non-categorical” to a “categorical” state in terms of eligibility to receive special education services. What this means is that, prior to this school year, students could receive services from schools without having to be given a diagnostic label such as “learning disabled” or “major depressive disorder.” Instead, decisions were made based on assessment results and clini-

Continued on page 24
SASP News

Within the realm of IQ testing, the impact of this change has been to make the WISC III and other cognitive measures of even greater importance when determining eligibility for services. Consider the case of learning disabilities, the most common diagnostic category within schools. In Massachusetts, the diagnosis of any learning disability is legally required to be based on the discrepancy model. That is, in order to be diagnosed as having a learning disability, a “severe” discrepancy between overall cognitive functioning as measured by a standardized test (typically the WISC III) and standardized scores in an achievement area must be established. Thus, if a child achieves a Full Scale IQ of 100 on the WISC III, using a 15 point differential criteria, a student would need a standardized score of 85 or lower on an achievement area to be considered “learning disabled.” If a child received a Full Scale IQ of 90 on the WISC III, the achievement cutoff also lowers, to 75. This creates a dilemma. The lower one’s IQ, the lower score one needs on achievement testing to receive academic support. The higher one’s IQ, the easier it is to be labeled “learning disabled” and thus receive support. This approach thus heavily favors those who score higher on IQ tests, who—surprise!—also disproportionately happen to be those that come from affluence and power within the community.

This situation places the school psychologist in a precarious position. Do you follow the law or the research? If one is involved in the assessment of dyslexia, for example, based on research one would likely want to assess whether or not the student has measurable difficulties with phonological processing—the primary deficit believed to be characteristic of dyslexia (see Siegal, 1999; Stanovich, 1999). The WISC III may measure many skills, but phonological processing is decidedly not one of these skills. If one follows the law and uses IQ scores as the primary determinant for a dyslexia diagnosis, one is arguably engaging in discrimina-
School Psychologists in Pediatric Settings: Intelligence Testing Considerations

Caroline G. McKnight
University of South Carolina
SASP News Columnist

Imagine sitting in a comfy child-size chair with your favorite test kit in front of you and the forever-cooperative child across the table. However, instead of the familiar chalkboard, book bags, and pencil sharpeners that usually surround you, there is a different scene. A few doors down, young children, who are medically fragile, play together, developmental pediatricians explain side effects of drugs to parents, and neurologists examine MRIs of a child’s brain to rule out possible effects of a seizure. This was the scene for me a year ago, as an intern at The Children’s Hospital in Greenville, South Carolina. As a wonderful experience, it provided a challenging alternative to previous practice in school settings. The movement of school psychologists from schools to alternative settings such as hospitals is growing. Differences in roles, responsibilities, and procedures allow for challenging experiences typically not encountered in the schools. Some of those differences are found in intelligence testing.

It may seem out of context to have a child’s intellectual functioning assessed in a medical setting; however, understanding the level of a child’s current cognitive ability can be valuable when explaining a diagnosis to a child, or when considering behavioral interventions. Within the pediatric setting, whether the child has been referred for an evaluation by an outside source, or is a patient in the hospital, there are general considerations when assessing his or her intelligence.

Initial Decisions

In the beginning stage of assessment, there are a variety of issues to consider when tests are being selected. Primarily, the purpose of the assessment should be determined. For example, awareness of the level of a child’s current intellectual functioning may be important before explaining an upcoming medical procedure. Or, the purpose may be to gain information about a child’s intellectual functioning following an operation. Furthermore, it is necessary to ascertain the appropriateness of conducting an evaluation. While a child’s IQ may seem important at the time, waiting until the child has regained energy, strength, and interest in working with someone new, may be the favored decision. It is of paramount importance to consider the child’s emotional and physical well-being first. A child may not be ready to sit for a full cognitive assessment and a screening instrument may more fitting.

Considerations During Assessment

When the assessment begins, there are procedural matters to bear in mind. Similar to testing a child in a school, adjustments need to be made so the child can be as comfortable as possible. However, when testing children who are patients in a hospital, there may be extra equipment not typically seen in a school. For example, the child may be encumbered by an IV, a gastrointestinal tube (G-tube), or may need a medical assistant close-by in the event of an emergency. Therefore, frequent breaks, comfortable seating, and other accommodations may need to be made. A significant component to the assessment is consultation with the child’s physician especially, and other medical personnel who have been involved in the child’s treatment and recovery. In the schools, information gained from multiple informants is ideal and a practice toward which to strive. In pediatric settings, where children may be suffering from chronic illnesses that greatly affect their academic performance, social and emotional well-being, it is critical to gather information from multiple sources. Often a transdisciplinary team approach is model to follow.

After the Assessment

Subsequent to assessment and consultation, report writing and communication of results usually commences. It is important to re-examine the purpose of the evaluation and how the results will affect the child’s overall well-being. The results may be a significant component of a transition plan as the child returns to school and home. Or, the physician may use the evaluation only as supplementary information. Thus, the goal of the assessment should be taken into account when communicating the results.

With the proper education of, and training and experience with children in pediatric settings, school psychologists can be important members of a team among the many individuals who contribute to the well-being of children with illnesses. Not only are school psychologists important in pediatric settings, but also they are an integral part to children’s transition from the hospital, back to the school. Most importantly, working with children with chronic or acute illnesses provides for many educational, meaningful, and memorable experiences. So, in spite of the absence of colorful bulletin boards, crayons and chalkboards, engaging a child with blocks, pictures and puzzles is just as entertaining.

In future segments of this column, other topics pertaining to the role of school psychologists in pediatric settings will be explored. For additional information, see School Psychology Review (1999), Vol. 28 (2): Mini-Series: Promoting School Success in Children with Chronic Medical Conditions.
Call for Nominations:
Editor-Elect, School Psychology Quarterly

Division 16 (School Psychology) of the American Psychological Association announces the opening of a search for Editor-Elect of School Psychology Quarterly. The Editor assumes a five-year, nonrenewable term to begin January, 2003 (a one-year Editor-elect term will begin January, 2002). The Editor should display a commitment to the science and practice of school psychology and embrace the standards and objectives of Division 16. Editorial experience; scholarly accomplishments; professional experience; administrative skills; and a commitment of time, energy, and some institutional resources are required. Interested individuals should submit a complete application packet, including:

1. A letter of intent, to include:
   • A statement regarding proposed scope and direction for the journal under a new editorial office
   • Personal beliefs regarding school psychological research and practice
   • Documentation of commitment to the editor position (approximately 1-2 days per week)
   • Commitment to APAs Ethical Principles and Specialty Guidelines
   • Statement and documentation of institutional support
   • Experience as a researcher, trainer, and practitioner
   • Editorial experience
   • Additional skills and experiences that can contribute to the role of the Editor

2. At least three letters of reference, to address perceptions of the applicant’s:
   • Skills as a writer and editor
   • Fairness, professionalism, integrity, and objectivity
   • Reliability, time management, and commitment to school psychology
   • Leadership in school psychology

3. Reprints of no more than three scholarly articles published in refereed journals

4. Current curriculum vitae

Applications should be submitted by May 1 to:
William P. Erchul, Ph.D.
Vice President of Publications, Communications, and Convention Affairs
Department of Psychology
North Carolina State University
Stinson Drive, 640 Poe Hall
Raleigh, NC 27695-7801
Phone: (919) 515-1709
Fax: (919) 515-1716
Email: william_ercul@ncsu.edu

ANNOUNCEMENTS

DIVISION ONE ANNOUNCES AWARD WINNERS

The Society for General Psychology, Division One of the American Psychological Association, announces its Year 2001 award winners who have been recognized for outstanding achievements in General Psychology. This year the winner of the William James Book Award is Michael Tomasello for his book The Cultural Origins of Human Cognition, which was published in 1999 by Harvard University Press. This award is for a recent book that serves to integrate material across psychological subfields or to provide coherence to the diverse subject matter of psychology.

The Year 2001 winner of the Ernest R. Hilgard Award for a Career Contribution to General Psychology is Murray Sidman. The winners of the George A. Miller Award for an Outstanding Recent Article in General Psychology are Jack Martin and Jeff Sugarman of Simon Fraser University for their article Psychology’s Reality Debate: A Levels of Reality Approach. The article appeared in the Journal of Theoretical and Philosophical Psychology in 1999 (pp. 177-194). Each award winner received a certificate and a cash prize: $500 for the Hilgard and Miller awards, and $1000 for the William James Book Award. The winner of the competition, to be determined and announced later, will deliver the Year 2001 Arthur W. Staats Lecture for Unifying Psychology and receive an award of $1000.

For all of these awards, the focus is on the quality of the contribution and the linkages made between the diverse fields of psychological theory and research. The Society for General Psychology encourages the integration of knowledge across the subfields of psychology and the incorporation of contributions from other disciplines. The Society is looking for creative synthesis, the building of novel conceptual approaches, and a reach for new, integrated wholes. A match between the goals of the Society and the nominated work or person will be an important evaluation criterion. The Staats Award has a unification theme, recognizing significant contributions of any kind that go beyond mere efforts at coherence and serve to develop psychology as a unified science. The Staats Lecture will deal with how the recipient’s work serves to unify psychology.

There are no restrictions on nominees, and self-nominations as well as nominations by others are encouraged for these awards. For the Hilgard Award and the Staats Award, nominators are asked to submit the candidate’s name and vitae along with a detailed statement indicating why the nominee is a worthy candidate for the award and supporting letters from others who endorse the nomination.

For the Miller Award, nominations should include: vitae of the author(s), four copies of the article being considered (which can be of any length but...
must be in print and have a post 1995 publication date), and a statement detailing the strength of the candidate's article as an outstanding contribution to General Psychology.

Nominations for the William James Award should include three copies of the book (dated post 1995 and available in print), the vitae of the author(s), and a one-page statement that explains the strengths of the submission as an integrative work and how it meets criteria established by the Society. Textbooks, analytic reviews, biographies, and examples of applications are generally discouraged.

Winners will be announced at the fall convention of the American Psychological Association the year of submission. Winners will be expected to give an invited address at the subsequent APA convention and also provide a copy of the award address for inclusion in the newsletter of the Society.

All nominations and supporting materials for each award must be received on or before April 15, 2001. Nominations and materials for all awards and requests for further information should be directed to General Psychology Awards, c/o C. Alan Boneau, Department of Psychology, George Mason University, Fairfax, VA, 22030. Phone: 301-320-3695; Fax: 301-320-2845; E-mail: aboneau@gmu.edu.

POSITION ANNOUNCEMENT NEW YORK UNIVERSITY

The School Psychology Programs of the Department of Applied Psychology invite applications for a senior position with a specialty in an area relevant to School Psychology. We seek candidates who have an established program of research, have a record of grant attainment, are scientist-practitioners, and who will contribute to our commitment to multiculturalism, diversity, and human development, in both research and practice, particularly as applied in an urban setting. This commitment extends to and characterizes the two APA-accredited doctoral (Ph.D. and Psy.D.) programs in school psychology and the certificate program in school psychology.

Responsibilities include graduate teaching, clinical and research supervision, engagement in an active research program, and student program advisement. This position provides an opportunity for program leadership.

Applicants must have a doctorate in school psychology or, alternatively, in child clinical psychology or community psychology, be eligible for licensure in New York State as a psychologist, have evidence of high quality research, scholarship, and grant obtaining, as well as a commitment to graduate and undergraduate teaching.

Salary: Competitive and commensurate with qualifications. Review of applicants will begin immediately and continue until the position is filled. Send letter of application, curriculum vitae, and sample(s) of scholarly work to: Professor Judie Alpert, Chair of School Psychology Faculty Search Committee, Department of Applied Psychology, New York University, School of Education, 239 Greene Street, 5th Floor, New York, NY 10003. Minorities are strongly encouraged to apply. NYU is an Equal Opportunity/Affirmative Action Employer.

PSYCHOLOGISTS AFFIRMING THEIR GAY, LESBIAN, AND BISEXUAL FAMILY

Organizing Committee: Laura S. Brown, Donald K. Freedheim, Marvin R. Goldfried, Ritch C. Savin-Williams, and Wendy K. Silverman.

This is to announce the beginnings of a newly formed network of psychologists who are coming out in open support of their gay, lesbian, and bisexual family members. As psychologists, we feel it is time to have our voices heard, especially in light of the hateful and damaging misinformation being put out by such people as “Dr. Laura” and those professionals advocating so-called conversion/reparative therapy. The goals of the network are as follows: 1) to openly support our GLB children, grandchildren, brothers, sisters, nieces, nephews, cousins, aunts, uncles, mothers, and fathers; 2) to impress upon less accepting family members the importance of supporting their GLB relatives; 3) to advocate gay-affirmative research and clinical work on GLB issues; 4) to encourage mainstream psychology to recognize and incorporate clinical and research work on GLB issues; and, 5) to assist GLB advocacy groups by providing research information and by testifying before legislative bodies.

At this stage, the network is new, but growing steadily. If you would like to add your name to the list, or would like to pass this information on to a colleague or family member, please e-mail, write or call:

Marvin R. Goldfried, Ph.D.
Professor of Psychology
State University of New York
Stony Brook, NY 11794-2500
(631) 632-7823
marvin.goldfried@sunysb.edu

ANTICIPATED POSITIONS AVAILABLE FAIRLEIGH DICKINSON UNIVERSITY

1) Director of M.A. Program in Clinical Psychology and Associate/Full Professor of Psychology (tenure-track, Spring or Summer 2001, FDU-Israel Campus). This position carries primary responsibility for administrative oversight of the School’s M.A. Program in Clinical Psychology in Tel Aviv, Israel and for teaching, research and supervision of students enrolled in the program. Requirements include an earned doctorate in clinical psychology from an APA-accredited program, a record of successful administrative and teaching experience in a similar or related position, a distinguished record of research and scholarly achievement commensurate with a faculty rank of senior associate or full professor, licensure or eligibility for licensure in New Jersey and/or New York, and familiarity with the training and practice of clinical psychology in Israel. Oral and written proficiency in Hebrew is desired, but not required. This position will require permanent residency in Israel beginning in the summer of 2001. In addition to a competitive salary and benefits package, a generous housing allowance is included for this position. 2) Director of Center for Psychological Services and Clinical Assistant/Associate Professor of Psychology (non-tenure-track, July 2001). This position carries primary responsibility for administrative oversight of the School of Psychology’s Center for Psychological Services in Hackensack, NJ, which serves as a training facility for the School’s doctoral programs in clinical and school psychology, and other special programs. Major responsibilities include coordinating/teaching clinical practica, clinical supervision, developing and marketing new and existing programs, facilitating research projects, grant writing, and maintaining positive relations with external agencies and the community-at-large. Requirements include an earned doctorate in clinical psychology, NJ license (or eligibility), and successful administrative and teaching experience. A record of scholarly...
achievement is also desired. 3) Assistant Professor in School Psychology (tenure-track, Fall 2001). This position is for teaching, research, and supervision of students in the School’s Psy.D. and M.A. Programs in School Psychology in Teaneck, NJ. Requirements include an earned doctorate in school psychology from a NASP-approved and/or APA-accredited school psychology program, and evidence of, or potential for, research and scholarly achievement commensurate with the rank of assistant professor. Experience as a practicing school psychologist and ability to contribute to the Ph.D. Program in Clinical Psychology (NJ- and/or NY-licensed or license-eligible) are also desired. The School of Psychology is located on the Teaneck-Hackensack Campus of Fairleigh Dickinson University in northern New Jersey, near New York City. In addition to offering graduate programs in school psychology, the School offers an undergraduate and graduate (master’s and doctoral) programs in New Jersey, the School offers an undergraduate and graduate (M.A.) program in clinical psychology at the University’s branch campus in Tel Aviv, Israel. Screening of applications will begin immediately and continue until all positions are filled. To apply for any of these positions, forward a cover letter, curriculum vitae, and the names, addresses and phone numbers of three references to: Dr. Christopher A. Capuano, Director, School of Psychology (T-WH1-01), Fairleigh Dickinson University, 1000 River Road, Teaneck, NJ 07666. Fairleigh Dickinson University is an Equal Opportunity/Affirmative Action Employer committed to a diversified workforce M/F/D/V.
The Division of Learning Disabilities of the Council for Exceptional Children (http://www.cec.sped.org/) provides up-to-date information on effective, research-based teaching strategies used by educators to help kids learn.

Here are links that could help you, help students develop/refine how they achieve, and at the same time, improve/expand their varied learning strategies:

1. http://www.TeachingLD.org  Teachers of students with learning disabilities;

2. www.muskingum.edu/~cal/database/encoding.html  Specific encoding and retrieval strategies.


Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Effect of Speed on Different WISC-III Subtest

(Perfect Score on subtest with no bonus points for speed)

<table>
<thead>
<tr>
<th>Age</th>
<th>PA</th>
<th>BD</th>
<th>OA</th>
</tr>
</thead>
<tbody>
<tr>
<td>6:0 to 6:3</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>6:4 to 6:7</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>6:8 to 6:11</td>
<td>18</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>7:0 to 7:3</td>
<td>17</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>7:4 to 7:7</td>
<td>16</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>7:8 to 7:11</td>
<td>15</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>8:0 to 8:3</td>
<td>14</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>8:4 to 8:7</td>
<td>13</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>8:8 to 8:11</td>
<td>12</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>9:0 to 9:3</td>
<td>12</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>9:4 to 9:7</td>
<td>11</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>9:8 to 9:11</td>
<td>11</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>10:0 to 10:3</td>
<td>10</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>10:4 to 10:7</td>
<td>10</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>10:8 to 10:11</td>
<td>9</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>11:0 to 11:3</td>
<td>9</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>11:4 to 11:7</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11:8 to 11:11</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>12:0 to 12:3</td>
<td>8</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>12:4 to 12:7</td>
<td>8</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>12:8 to 12:11</td>
<td>8</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>13:0 to 13:3</td>
<td>7</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>13:4 to 13:7</td>
<td>7</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>13:8 to 13:11</td>
<td>7</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>14:0 to 14:3</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>14:4 to 14:7</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>
Effect of Speed on Different WISC

<table>
<thead>
<tr>
<th>Time Range</th>
<th>14:8 to 14:11</th>
<th>15:0 to 15:3</th>
<th>15:4 to 15:7</th>
<th>15:8 to 15:11</th>
<th>16:0 to 16:3</th>
<th>16:4 to 16:7</th>
<th>16:8 to 16:11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Values</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

Adapted from WISC-III Manual

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
In preparation for the upcoming reauthorization of IDEA, the US Office of Special Education Programs (OSEP) has commissioned leading researchers in the field of Learning Disabilities to write major papers on nine topics:

* Learning Disabilities: Historical Perspectives
  

* Is Learning Disabilities Just a Fancy Term for Low Achievement?
  

* Judgments in Identifying and Teaching Children with Language-based Reading Difficulties
  

* Early Identification and Interventions for Young Children with Reading/Language Disabilities
  

* Empirical and Theoretical Support for Direct Diagnosis for Learning Disabilities by Assessment of Intrinsic Processing Weaknesses
  
* Classification of Learning Disabilities

http://www.air-dc.org/ldsummit/download/Fletcher%20Final%208-10-01.pdf

* Learning Disabilities as Operationally Defined by Schools


* Discrepancy Models


* Responsiveness to Interventions: An Alternative Approach to the Identification of Learning Disabilities


These papers are available (as PDF files) in their entirety at the links above.

Carlo Cuccaro

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
With all the talk on the list about LD and the reauthorization, I though some might get a kick out of a handout I picked up at a conference on “How to help school aged children” subtitled “How to sue a school.” This conference was attended by about 200 attorneys and was a primer on how to get kids into special education.

I will not reveal who gave the handout, but it was a representative of a well known national Learning Disabilities Advocacy group. It is given verbatim with only my random comments interspersed

Is it any wonder that schools have problems? If this were not real it would be hilarious.

**Basic Knowledge of Learning Disabilities**

**What is a Learning Disability?**

There are over thirty five accepted definitions of learning disabilities--from the federal
government, education system, state systems of education, health and assistance. *Primarily, a learning disability is when there is a severe discrepancy between the expected IQ and the actual IQ. The client is working at an achieved IQ much lower than is proven possible.*

(underline, italic added – comment – those ld kids sure do the impossible)

**Effects on Family Life with Learning Disabled Member**

Parental reaction to LD diagnosis is more severe than any other disability. *(comment – have you ever had to explain to the parent of an autistic child or one with mental deficiencies the results and implications of the comprehensive evaluation? I'll take the LD child almost anytime)* it is hidden and takes usually failure and loss of self-esteem to even have testing done such a disability. Most other disabilities can be seen from the beginning. LD is found out later, after dreams and aspirations have been contemplated about that child. Suddenly the future and dreams are blown away with two words--“Special Education”. Unfortunately, most people are very uninformed about learning disabilities *(comment – as evidenced by this handout)* and immediately feel that this means mental retardation. Friends of the family will say, "Gee, I'm so sorry, he doesn't look different." All of this can tear a family apart. In fact *(comment - this is what I want – facts)* it has been proven that most disabilities will bring the family closer together whereas learning disabilities often tear the family apart. In fact *(comment – more facts)* it has also been proven that many fathers do not do well with learning disabilities--usually sons (7 out of 10, are males). *(comment – 7 out of 10 fathers are male? Or 7 out of 10 sons are male?)* Parents who agree usually agree on everything, will not
IF only this were made up

agree on the LD

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Learning Strategies Guides

Following are links to study guides that are available on-line. Many of these are maintained by the academic services centers at colleges and universities, but much of their information is useful for high school and middle school students, as well. Each site features a variety of study guides, covering a wide assortment of topics, and usually about one to three pages in length. Many of the same topics are covered at each site, but each may have its own unique style or emphasis that may be most helpful for you. Find the study guides which you like best, or which suit your style of teaching or learning, print them out or download them.

High School Hub
Academic Resources for High School Students
http://www.highschoolhub.org/hub/
The High School Hub is a noncommercial gateway to excellent free online academic resources for high school students. It features a study guides, reference resources, college information, and subject guides for English, mathematics, social studies, science, global languages, arts, and technology, and much more.

Study Skill Guides
Dartmouth College, New Hampshire
http://www.dartmouth.edu/admin/acskills/index.html#study
Study Strategies Homepage
University of Minnesota-Duluth
http://www.d.umn.edu/student/loon/acad/strat/
Learning Skills Center
University of Texas at Austin
http://www.utexas.edu/student/lsc/makinggrade/makinggradeindex.html

Study Skills Guides
College of Saint Benedict & Saint John's University, Minnesota
http://www.csbsju.edu/academicadvising/helplist.html
Basic math and algebra for the community college adult learner,
Camden County College, Blackwood, New Jersey
Instructional Support Services/ Learning Center
St. Thomas University
St. Paul, MN
This site has extensive resources of its own, as well as links to many other sites.

Academic Skills Center, California Polytech
http://sas.calpoly.edu/asc/ssl.html

Improve Your Studying Skills
University of North Carolina at Chapel Hill
www.unc.edu/depts/unc_caps/TenTraps.html

General-Purpose Learning Strategies
Muskingum College New Concord, OH
www.muskingum.edu/~cal/database/genpurpose.html

Excellent and extensive guides to studying and learning, and explanations of how the learning process works.

The CalREN Project Study Tips University of California at Berkeley
http://slc.berkeley.edu/CalRENHP.html

Study Skills Self-help Information
VA Polytechnic Institute and State University
www.ucc.vt.edu/stdysk/stdyhp.html

An annotated bibliography of study skills books by Arlene Young, Student Services
Athabasca University, Athabasca, Alberta, Canada
www.athabascau.ca/html/services/advise/ssbib.htm#sec3

Study Skills Package
University of Waterloo, Ontario, Canada
www.adm.uwaterloo.ca/infocs/Study/studyindex.html

Basics of Reading:
American Studies Learning skills
www.dla.utexas.edu/depts/ams/lskills/

eStudyCentre, from Allen & UNWIN Independent Book Publishers, Australia, is designed for university, college and TAFE students in all disciplines and at all levels, whether you’ve just started first year or are working on a postgraduate degree.

Prepared by Jonas Taub, School Psychologist, Milford High School

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
One way of "sorting out" cognitive abilities is to try to create factorially "pure" tests. Wechsler never intended his subtests to be used this way, and they cannot, but we can paradoxically take advantage of shared demands among the very mixed measures of the Wechsler scales as an alternative approach to the same goal.
The Lindamood Program for Reading Spelling and Speech is excellent.

http://www.wilsonacademy.com/ Wilson Reading program. Wilson is used with all age groups and has built in curriculum based assessment.

http://www.projectread.com/ Project Read is worth looking at if the district is serious about prevention:

http://at-advocacy.phillynews.com/read/reading2.html A good overview of a variety of synthetic-phonetic programs

http://www.sciam.com/1196issue/1196shaywitz.html Overview of current research on dyslexia written for the layman,

http://www.interdys.org/ The International Dyslexia Society. They will do a zip code search to help you find experts who have been trained in your area.
I wrote this as material to give parents:

Attention in School: Some Advice

If you are reading this, you probably want to find a way to increase your child’s ability to pay attention and concentrate in school without having him or her take medication. There are so many reasons that a children (or adult) does not pay attention that it is impossible to list them all. Saying that there is one reason for a child’s activity level or inattention in school is a lot like saying that we know a factual reason for everything that happens in the world.

The first step before taking a journey involves getting organized. Get a notebook or a journal and two pens or pencils. You are going to be taking a bunch of notes. The next step involves making yourself comfortable so that you can read this. Your child’s problems probably began with a teacher saying that he (most problems with attention happen to boys) or she is moving around too much, or not paying attention. The first place to begin is to rule out physical/medical problems. Has your child had a thorough physical, including vision testing (not just, "Look over there and read that."), hearing (some physicians whisper to a child to test hearing – that is never enough), blood tests (thyroid problems, for example, can cause inattention, as can problems with blood sugar, and other things), neurological screening, etc? I advise that you tape record your interactions with your primary care physician concerning the attention problem, and offer to give him or her a copy of the tape.

The second place to look at is your child’s school. Have you spent time in his/her classrooms? While embarrassing to middle school and high school students, this is still an important thing for a parent to do. Look around the classroom. Think about whether the physical layout is good for learning. Write your observations in your notebook. Watch the teacher as s/he teaches. Ask yourself whether you could learn from this person. Note how many students ask questions, and how they are answered by the teacher. Note (in your book) how many times s/he asks students questions. Note how many times s/he asks students who do not raise their hand questions. Look around and find those students who appear to have problems paying attention (they might be talking to another student, playing with a pencil, etc.). Note how many times the teacher asks these students a question about the material being taught.

Yes, visiting the school involves taking time off from work. In some states (North Carolina is one of them), there are laws that require employers to give employees time off to deal with school matters.

Talk to your child about school every day. Asking a question like, "What did you do in school today?" is a sure way to get discouraged because most children beyond second or third grade know that you will leave them alone if they say things like,
"Nothing," "The usual," "Stuff," etc. as an answer. You must be specific and persistent. Get to know the names of you child’s teacher(s). Say, "Tell me two (three, four, etc.) questions that you asked Ms. Smith today?" Be prepared to tell your child that s/he needs to write these down along with the teacher’s responses every day, because you are going to ask. Talk with your child about his/her response. Write this in your journal. Ask your child to write down every time that the teacher asks him/her a question about a lesson. Make certain that he writes his/her response. Write this in your journal.

Some children are going to resist talking to you. Persist. Make daily conversations about school as important as brushing teeth. This has got to become a habit. Make it more pleasant by having the conversations while eating a healthy snack.

This is a lot of work for you. Take it to the person who works with you to discuss. Are you satisfied with the number of times that the teacher is talking to/questioning your child each day? Discuss ways to make this better.

The Job Jar is an added way to get your child to talk about school. It needs to be used in addition to the daily school conversation, so try the daily conversation first. Generally, the more you pay attention to what your child is doing in school, the more your child will pay attention in school. Continue to work on this.

**Working with School**

Children with significant attention problems are often referred to special committees for assistance or support. Sometimes a committee recommends an evaluation that results in a special education placement. Sometimes a child may have a "diagnosis" that makes him or her eligible for accommodations under the Civil Rights Act. In any case, there is a chance that goals will be established for your child. You have the right to participate in the creation of those goals.

When a child has difficulty paying attention, the most important goals are those which enhance it. One goal must be, for each class, "________(name of student) will demonstrate that they have been paying attention to the class by successfully responding to one or more questions about the topic. S/he will be asked at least one question per class regardless of whether a hand is raised to volunteer a response. Every effort will be made to ask a question at those times when the student is most distracted (for example, when talking to another student, or when fidgeting)."

The student must be chosen to be a leader in group project situations in order to enhance concentration and benefit from peer support. These leadership positions must be assigned periodically and randomly when possible.

Teachers can enhance attention in class by providing breaks. Breaks can consist of: 1) A Daydream Break, in which students will be given 1 to 2 minutes to close their eyes and think of a preferred activity or place to be. The teacher can provide guided imagery, or ask a student to do it; 2) The teacher can ask the class to stand-up and stretch.

Students need activity breaks during the school day. Recess should be divided so as to provide at least one break in the morning in elementary school and one break in the afternoon. Silent lunch is a poor form of punishment because it deprives students of an activity outlet. Deprivation of recess is a poor form of punishment because it also deprives students of activity that they need in
order to pay attention successfully. Some forms of motor activity that violate minor rules can reduce fidgeting in some circumstances. For example, allowing a student to chew gum can sometimes reduce fidgeting.

One rarely accessed form of building movement into class structure is the response format. Students can be required to stand each time they ask the teacher a question, or when they respond to a question. This form of movement is beneficial for three primary reasons: 1) It wakes up students who may be drowsy; 2) It reduces discomfort from remaining seated for long periods; and 3) It reduces acoustical problems that may come from sitting. The acoustical consideration is rarely mentioned. Think of what happens when sound strikes absorbent material. It stops or is deadened. Speaking while seated at the back of a class, for example, requires that the sound project through the seated bodies and clothing of students forward of the position.

Students in middle school and high school often get into trouble during the change of classes because those change periods are really breaks in the tedium of sitting for long periods. Careful supervision of these transitions can allow for the "break" to be natural, without being a source of difficulty.

Students who attend "Blocked Classes" (classes lasting for one hour and fifteen minutes or longer), need breaks. These are the worst kind of classes for children who have trouble paying attention. Parents are advised to send their children to schools with shorter classes or work with teachers to provide breaks.

Some students find that a brief meditation break can help them to establish self-control. Some students benefit from being given tasks to perform that interfere with transition problems (such as carrying a "note" from their teacher to the teacher of their next class).

**The Job Jar**

The Job Jar is an idea that is deceptively simple. Its purpose is to ask students questions at home that increase attention at school.

Materials needed:

1. Two large jars (preferably unbreakable)
2. A single standard game die
3. Small sheets of paper with tasks or questions

The child rolls the die. The number on the die determines how many pieces of paper/questions s/he takes out of the jar to answer.

Some pieces of paper contain the word, "PASS." This means that s/he can skip that question. Do not put the drawn pass back into the jar until all turns are taken. The child must respond or discuss what is on each sheet. You can create more questions as needed. The ones provided are just samples. Orient the questions to your child’s school and classes.

**Potential Problems**
Changing a child’s behavior is **NEVER** an easy task. Often a child’s behavior will get worse before it gets better. Here is where persistence pays off. Do not give up or say that something is unsuccessful until you have given it a very long try. Discuss problems with a neutral party, such as a counselor or a teacher.

### Memory Questions

Teach me something that you learned in your math class today.
Teach me something that you learned in your History/Social Studies class today.
Teach me something that you learned in your Communications class today.
Teach me something that you learned in your Science class today.
Tell me what question your Math teacher asked you today in class. What was your answer?
Tell me what question your Science teacher ask you today in class. What was your answer?
Tell me what question your Communications teacher asked you today. What was your answer?
Tell me what question your History/Social Studies teacher asked you today. What was your answer?
Tell me one question that a student asked in Math today. What was your teacher’s answer?
Tell me one question that a student asked in Science class today. What was your teacher’s answer?
Tell me one question that a student asked in Communications class today. The answer?
Tell me a question asked by a student in History/Social Studies today. The answer?
Tell one thing you learned in Math and one thing you learned in Science today.
Tell one thing you learned in Math and one thing you learned in English today.
What did your Math teacher wear today?
What did your Science teacher wear today?
What did your Communications teacher wear today?
What did your History/Social Studies teacher wear today?
Tell one thing that your Math teacher wrote on the board or overhead projector today.
Tell one thing that your Science teacher wrote on the board or overhead projector today.
Tell one thing that your History teacher wrote on the board or overhead projector today.
Tell one thing that your History/Social Studies teacher wrote on the board or overhead today.

### Softies

**PASS**
Talk about what you did in Physical Education today.
Talk about your lunch today.
What part of school was the most fun today?
**PASS**
What did you do/talk about with your friends during the change of classes today?
Talk about ways to make studying easier for yourself.
Talk about your easiest subject. Why is it easy for you?
If you had a magic power to change something about your school what would it be?
PASS
If you had a magic power to change your teachers into animals talk about it.
What class would you like to teach and why?
Name a student that you think would make a good teacher. Why?

Thinkers

Read the newspaper. How does what you have learned in school today relate to an article?
Talk about how something that you have learned in math can help with shopping.
Talk about how something that you have learned in history/social studies is related to current news.
Talk about how something that you learned in science is related to your life.
Is it easier for you to learn in the morning or the afternoon? Talk about it.
Talk about how school has helped you with something that you have found useful.
Look in the job advertising in the newspaper and find the jobs that don’t require a high school diploma.
Multiply the current minimum wage by 40 hours. Subtract taxes that you pay (get parents help). Use the newspaper to figure out a monthly budget.
If everyone received assistance from the government because they did not graduate from school, who would pay for it?
How did farmers learn to grow vegetables and raise animals? What kind of classes helped them?
How did people learn to forecast the weather? What kind of classes helped them?
How did cooks learn to cook? What kind of classes helped them?
What kind of things do you need to know in order to become an actor? How did school help them?
What kind of classes might help you to fill out an application for a job?
What kind of things that you learned in school might help you to take the written part of the test to get a driver’s license?
John Willis answered a query about the use of the Bender Gestalt's "Emotional Indicators." He noted at that time that although there was a significant Chi-square statistic, he wondered about the "odd way to discover" these behaviors. Below is John's response and an extension of John's concern. Using Positive and Negative Predictive Power statistics, more information about the Bender is revealed.

Elizabeth's Koppitz is sticking with her original assertion in the 1975 manual.

<table>
<thead>
<tr>
<th>Emotional Indicators</th>
<th>Good Bender</th>
<th>Poor Bender</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 2</td>
<td>119</td>
<td>76</td>
</tr>
<tr>
<td>3 to 6</td>
<td>19</td>
<td>58</td>
</tr>
</tbody>
</table>

Chi-square was p < .01 at ages 8 to 10 and p < .001 at ages 5 to 7. "This supports the hypothesis that children with poor visual-motor perception also tend to have a high incidence of emotional indicators on their Bender records. Three out of four of all subjects with three or more emotional indicators were also found to have a poor Bender score. However, there were subjects with good Bender scores who showed a high incidence of emotional indicators while more than half of all children with poor Bender records revealed less [sic] than three emotional indicators. Thus it seems safe to assume that the Developmental score and the emotional indicators on the Bender Test measure different aspects of a child's functioning but that both are found more often together on the records of children with emotional problems than on those of children without emotional problems. A poor Bender score alone, does not necessarily imply emotional problems, but if a child shows several emotional indicators as well as a poor Bender score, then indications are that the child has serious emotional problems and that perceptual problems have probably contributed to his disturbance" (p. 142).

This still strikes me as an odd way to discover such behaviors as impulsivity, acting out, timidity, and shyness.

Using parameters about a test's diagnostic usefulness that have been proposed by Elwood (1993), and using information taken from Elizabeth Koppitz's *The Bender Gestalt Test for Young Children*, page 130, the following is found:

Test specific parameters include sensitivity, or the proportion of individuals with a disorder that exhibit the sign (i.e., the proportion of children with Emotional Problems who receive scores within the abnormal range - in this case 43.4%) and specificity, or the proportion of individuals without a disorder that do not exhibit the sign (i.e., the proportion of controls who receive scores within the normal range - in this
BENDER GESTALT and EMOTIONAL PROBLEMS (case 86.8%).

<table>
<thead>
<tr>
<th>EP Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;2</td>
<td>59</td>
</tr>
<tr>
<td>&lt;3</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>136</td>
</tr>
</tbody>
</table>

Sensitivity = 43.4%  Sensitivity = \( \frac{a}{a+c} \)

Specificity = 86.8%  Specificity = \( \frac{d}{b+d} \)

These two parameters are calculated in the research setting by first knowing the diagnosis of the children (through test-independent criteria) and noting how they perform on the test of interest. However, as Ellwood (1993) points out, this is the opposite of the way an evaluator uses a test. The evaluator starts with the test score and attempts to determine the child's diagnosis. In order to judge the usefulness of a test for this purpose, the evaluator will need to look at a test's sensitivity and specificity in light of the disorder's base rate in their referral population.

For example, if a Bender was used as a screening measure on a population of 1000 children in which 10% (100) of the children have “Emotional Problems”, and that test gives an abnormal score for 43% of the children with “Emotional Problems” (i.e., sensitivity) and gives a normal score for 87% of the children without “Emotional Problems” (specificity), the following diagnostic properties result.

| EP Control |
|------------|---|
| >2         | 43 | 32 | 75 |
| <3         | 57 | 868 | 925 |
|            | 100 | 900 | 1000 |

PPP = 57.3%  PPP = \( \frac{a}{a+b} \)

NPP = 93.8%  NPP = \( \frac{d}{d+c} \)

Using this table, one can calculate Positive Predictive Power (PPP), or the chances that a child who receives an abnormal test score actually
BENDER GESTALT and EMOTIONAL PROBLEMS

has “Emotional Problems”. PPP = a/a+b = 43/75 = 0.57. A test with 43.4% sensitivity and 86.8% specificity has restricted usefulness as a diagnostic tool if it is used on a population with a 10% base rate of the disorder because if the child receives an abnormal score, (s)he is almost as likely to be a control (42.6%) than a child with “Emotional Problems” (57.3%)


Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
A question has come up about how to test kids who have been retained. On the Woodcock, do we score them by age, or by grade?? It would seem to me that they should be scored by grade, since it would not be fair to score them with their peers in a higher grade since they have not been exposed to the material or subject matter at the higher grade level. Let me know. Thanks! Kim

In my experience with younger kids, there is little statistical difference in the standard scores when scoring by age or by grade in the case of retention. I typically will score a retention case both ways on the scoring program and then compare them. They tend to meet the discrepancy either way. This may be different with older kids, but if they are run through the computer both ways, the examiner will be able to make very clear judgments regarding what is instructionally related and what is learner related.

Subject: RE: Grade/Age Equivalents

You are most correct in that many different interpretations of age/grade equivalents are made and most of them are WRONG. It is for this reason and the fact that they are not equal interval metrics that measurement experts for the past 80 or so years have warned against their use. The usual meaning that parents and teachers come away with when grade/age equivalents are used is that the child has the skills approximately of the age/grade specified which in not at all the case. The age/grade equivalent is simply that the child obtained the raw score corresponding to the average raw score obtained by a group of children at the specified age/grade. Most tests lack a sufficient number of items at any specific age/grade level of difficulty due to their broad age/grade range of the test. I recommend to my undergraduate and graduate students to present only standard scores and percentiles and interpret those. Age/grade equivalents are so misused that they create more problems than answers. Sadly, I have heard psychologists misinterpreting age/grade equivalents to parents and teachers!

Gary

Subject: Re: Grade/Age Equivalents

A further caution about the age equivalents on the WISC-III. John Willis and I did a study examining the intra-subtest scatter on the subtests of the WISC-III. (Dumont, R. & Willis, J. O. (1995) Intra-Subtest Scatter on the WISC-III for Various Clinical Samples vs. the Standardization Sample: An examination of WISC Folklore?)
What we found was that it was quite normal for someone to have scatter within a subtest. Since the age/grade equivalents are interpolated from the raw score, the fact that a person can have quite a bit of scatter within the scale suggests that the age/grade equivalents are further confounded by normal variation. A person can have a raw score of 10 on information by answering the first 10 items correctly and failing all others, or get a raw score of 10 by answering correctly every other item from 1 to 20. Same raw score, very different performance (but remember scatter is normal).

Hope this is useful.

Ron

I am only familiar with a table of age equivalents that appears in the WISC-III supplementary tables in the back of the Manual; I use the other Wechslers so seldom in my current assignments I'm not sure what they have. The WISC-III gives age-equivalents for subtest raw scores; and my interpretation is that this raw score was the average score obtained by students of the given age. At least that's usually the implication. I don't know how this specific table was generated, but often in drawing up age-equivalent tables the means at specific ages are plotted, and ages in between obtained means are interpolated. My understanding is an overall developmental age for the WISC-III may be estimated by computing a median among the obtained subtest age-equivalents.

I use WISC-III age equivalents very cautiously, usually only where it may be useful to have an approximate cognitive development level. For example, a child referred as possibly having AD/HD, but who is cognitively delayed, is going to be developmentally like a much younger child in the degree of distractibility and impulsivity they demonstrate. If they demonstrate distractibility and impulsivity relative to same-age peers, it may be due to general cognitive delays, rather than AD/HD. There are times when a developmental age is useful to emphasize a student's ability to reason and use good judgement; e.g., in a manifestation determination hearing on an MR student who was caught with drugs -- it may be useful for the Multi-Disciplinary Team to understand that this is a 14 year old who reasons like an 8 year old.

Having said all that, age and grade-equivalents must carry a lot of caveats. They are not equal-interval scales (going from age 8 to 9 is actually a greater amount of growth than going from age 11 to 12), and because of that, you cannot apply ordinary arithmetic (which is the reason for using median age-equivalent as a measure of central tendency), as you might want to do in showing growth over time. Variability in scores is not standard throughout the range of the scale, so a "two year deficit" is not as severe at some ages as it may be at others. Grade- and age-equivalents appear to be more sample-sensitive than standard scores; that is, there appears to be more variability between tests in grade- and age-equivalents, and more consistency when comparing standard scores between tests, when samples are comparable but don't contain the same subjects.

John
Gary, partly your sadness is due to the failure of training programs to provide a thorough grounding in the relationships between data types (categorical, interval, etc.) and the metrics used to describe them, and partly also a failure of the interpreters' to understand that the underlying distributions in skills represented by different grade equivalents can be enormously different; the reading skills difference, for example, between a beginning 2nd and 3rd grader is much larger than that between a 9th and 10th grader, but both differences "look" like 1 year.

Frank

Sadly, I have heard psychologists misinterpreting age/grade equivalents to parents and teachers!

Subject: Re: Grade/Age Equivalents

Some years ago, the International Reading Association condemned the use of grade equivalents for reading tests. Good for them! Reading Teacher, January 1982, p. 464.

John Willis

Subject: Re: Grade/Age Equivalents

I had the pleasure of testifying in a due process hearing in which a doctoral-level neuropsychologist swore under oath that a student was working at a fourth grade level in arithmetic (the student's grade level on the WRAT). Since the student was in eighth grade, that seemed very sad. However, the subpoenaed protocol, showed the student computing with common fractions with unlike denominators, doing long division with decimal fractions, and generally handling very advanced arithmetic competently. The low grade equivalent stemmed from simple errors on sums and products and occasional misreadings of digits and operations signs. The neuropsychologist not only misinterpreted her own findings, but convinced him/herself of the validity of the misinterpretation. Happily, the hearing officer was wiser.

John

Subject: Re: Grade/Age Equivalents

The complaints you've seen about age equivalents on the WISC also pertain to both age and grade equivalents on the WIAT. It doesn't matter how they are derived. Don't use them. Grade-level designations of items are a different matter. For example, reading inventories and some reading tests, e.g., Diagnostic Assessments of Reading with Trial Teaching Strategies and the Diagnostic Reading Scales by Spache, have the student attempt tasks designated at specified grade levels, so you can say Ralph handled the third grade material easily, the fourth grade material marginally, and crashed and burned on the fifth grade material.

John Willis
Date: 3/20/99 10:16pm

Subject: Re: [Fwd: age vs. grade norms -Forwarded]

Depends on what you are trying to prove. In the state of Colorado age norms are preferred. I generally will run a profile of both if there has been a retention because when there is a retention involved, the kid will always come out looking bad on age norms when it is actually a lack of opportunity for instruction due to the retention. If you are simply interested if the grade placement is appropriate for that child, then grade norms might be preferred. Otherwise, compare the child to others the same age. Should you run both age and grade norms on a child who has not been retained, you will find there is little to no difference in the standard scores.

Subject: Re: No Subject


John Willis

[NOTE: email the IRA and they may send you a copy. That's how I got one. - rogerp]

Subject: Re: Test Interpretation

Just a point about intrasubtest scatter (when a person fails "easy" items but goes on to pass "harder" items). John Willis and I investigated this phenomenon by sampling 410 students having various classifications (Id, adhd, s/l, premature birth). [ Dumont, R. & Willis, J. O. (1995) Intra-Subtest Scatter on the WISC-III for Various Clinical Samples vs. the Standardization Sample: An examination of WISC Folklore? Journal of Psychoeducational Assessment , 13, 271-285.] In our conclusions we wrote:

"Although intrasubtest scatter might be indicative of some dysfunction in adults, extreme caution must be taken when making the same interpretation with children. The face-value assumption that groups of children identified as needing some form of educational service would display more intrasubtest scatter than normal was not supported by this study. In fact, of the 12 cases where
scatter was evident and significantly different from the mean of the standardization group, 7 indicated significantly less scatter. Further analysis with subgroups based on both diagnosed disability and test age in fact demonstrated no meaningful difference among students with diagnosed disabilities, students referred and found not to have educational disabilities, and students in the original WISC-III norming sample. As seen from this study, the data on intrasubtest scatter suggest that it may be fairly common in those suspected of having learning disabilities, but the intrasubtest scatter was neither large nor consistent enough to serve as a diagnostic marker.

We called this sort of issue "WISC Folklore" since it sounded like it should make sense but when examined empirically, it was not supported.

Ron

Jim wrote:

> Sattler's book does a good job of explaining inter and intra subtest scatter. If I remember correctly, > if the range of subtests in the verbal or performance realm is 7 or great (e.g. Vocabulary 12 and > Information 4) intrasubtest scatter exists and must be explained in the interpretation section that > the result may not be an accurate representation of the child's true ability due to the scatter in > performance. Jim Lloyd

This article comes from:

CEC: TODAY: Exclusively for membership of the Council for Exceptional Children, Vol. 5 No. 4 November 1998

Assessments Fail to Give Teachers Relevant Information

Through students with disabilities are routinely given assessments to determine their eligibility for special education services, these assessments rarely provide special or general education teachers the information they need to provide effective instructions. Too often, teachers are given only an overall score for a student's IQ or achievement level, which provides only limited information about a student's functioning. In other cases, the assessments are too narrow to provide an accurate picture of the student's abilities.

To redress this situation, assessments need to be restructured and teachers need to receive more complete information, according to experts in the field. Rather than giving students a standard set of tests, such as Wechsler and Woodcock-Johnson, as is often the case, educational diagnosticians must broaden their use of assessment tools. In addition, educational diagnosticians should know when they can appropriately provide students with accommodations for individual assessments.

While assessments have always played a central role in determining a student's eligibility for special education, they have come under fire in the past few years. Critics claim that current assessment techniques do not accurately identify the presence of disabilities, particularly learning disabilities or behavioral disorders. Others say that through our assessments we focus too much on a student's weaknesses rather than his or her strengths. The need for refinement of assessment strategies has been further highlighted by the Individuals with Disabilities Education Act (IDEA) 1997, as well as current reform movements emphasizing standards. Because students with disabilities must have access to the general education curriculum and they are expected to meet the same high standards, teachers must have more information about a student's abilities and current functioning if they are to work with him or her effectively. "IDEA '97 points out the need for individual comprehensive evaluations," said Douglas Smith, professor at the University of Wisconsin, Fiver Falls.

Weaknesses of Current Assessment Practice
The primary problem with current assessment practice appears to be the lack of comprehensive information, either because the tests fail to provide it or the assessor limits the information that could be gleaned from the process. He or she may restrict the number of tests given or the number of subtests within an assessment tool.

Ironically, computerization has contributed to the lack of relevant information about a student's functioning that is passed on to the teacher. Rather than receiving the educational diagnostician's analysis of student achievement, learning patterns, and other information that can be used to determine instructional strategies, the teacher often receives an overall computerized score, according to Rosalind Rothman, professor at the Center of Sunny and Cuny, NY. Such scores might say a student is working at the third grade level but do not give any type of analysis of the student's abilities or problems.

Following is an overview of some of the assessments often used with students with disabilities and their strengths and weaknesses.

**Intelligence Tests**

The way we currently test for cognitive processing shortchanges teachers in two ways: we can derive a falsely high score and we do not learn in what ways learning is being interfered with, according to Gary Hessler, consultant for Malcomb Intermediate School District in Clinton Township, Mich. That is because examiners often administer only the first three subtests, which assess fluid intelligence (abstract thinking, problem solving), verbal intelligence, and non-verbal intelligence. Unfortunately, it is often the lower level cognitive processing that causes problems for our students, Hessler says. Lower level processing includes:

- **Long-term retrieval** - The ability to retrieve information on demand.
- **Short-term memory** - The ability to hold information in one's immediate awareness long enough to think about it.
- **Working memory** - The ability to remember information long enough to think about it and use the information to solve a problem.
- **Processing speed or automaticity** - How rapidly and automatically one can perform simple tasks (affects routine abilities like sight word knowledge and math facts).
- **Phonological awareness** - How well one understands that words are made up of sounds.
- **Orthographic ability** - How well one perceives and retains visual letter patterns. Fine motor ability - The ability to rapidly perform fine motor tasks, such as handwriting.

To use an intelligence test, such as the Wechsler, to its best advantage, we must go beyond an overall IQ score, recommends Hessler. Examiners should administer several of the subtests, evaluate each separate subtest, and get a sampling of each. From this information, the evaluator can see patterns and discern where the student is having a problem in cognitive processing. While all 10 of the intelligence factors should be considered, the student may not need to be evaluated formally in each area, says Hessler. Examiners can learn a lot about a student's capabilities by interviewing his or her teachers, parents, and others who work with the student.

**Achievement Tests**

While the Woodcock-Johnson Achievement Test (WJ) and the Wechsler Individual Achievement Test (WIAT) are two of the most commonly used assessments for students with disabilities, a study by Rothman showed that the tests have several weaknesses.

Reading. The WJ passage comprehension sections were viewed to be unrelated to what is done in the classroom. The WIAT subtest questions are poorly constructed...
and the passages are too short for older students.

Math. Neither the WIAT nor the WJ provided enough examples of different problem types in the computational math subtest. Again, the WIAT was inappropriate for older students. The problem solving section of the WJ relies too much of time and money to give valid assessment of student's abilities.

Written/Oral Language. Interpretation of written and oral language on both tests is very subjective. In addition, the WIAT's subtest on oral language should be redesigned. The WJ's written language subtest for the older grades is unrelated to the type of writing students are expected to perform.

While no one denies the need for testing, we must re-evaluate the instruments we currently use to make the assessments more relevant to what is expected in the classroom. The question is whether or not we can do that with our current instruments, Rothman says.

Marci, in one sense I guess it depends on what the reason is for reporting the scores. If you want to see how the child is doing compared to his grade peers, some people use the grade scores. The problem with using grade scores, however, is best seen through a little thought experiment. Let's assume a hypothetical first grade kid is expected to master adding a column of 3 single digit numbers and let's further assume our target kid has just learned to accomplish that task reliably. On the surface it sounds like he gets an "average grade standard score" for first grade on this task. Now what if I tell you the kid is 14. Hmmm. My point is that using grade scores for kids who have been retained (or as a measurement of any child of an arbitrarily older age than his grade peers) covers up the fact that our hypothetical kid is probably better thought of as having seriously deficient math skills rather than as having solid first grade skills. My experience is that grade standard scores are often used to give an illusory sense of accomplishment and justify the retention of kids through statements like "Johnny's doing so much better now that he has been retained. His scores are average!" There is often the additional tacit assumption that because Johnny's skills are first grade level, he can function in the first grade curriculum. That's often wrong too: Johnny's functioning is probably very different from a first grader, although he happened to get the same number of problems correct on the test. - This problem again came into focus recently for me when assessing an 8th grade student. Her scores, for the most part, were close to average for 8th grade students, but she was still failing all her core courses miserably. The tests said she shouldn't be having such serious problems when grade norms were used. However, the girl had been retained twice; she was actually two years older than her peers. Using age standard scores it was easy to make a case that she needed special services. This is not the first time I have seen students experience de facto denial of services because they had been retained. - The moral of the story now that I get off the soap box? Stick with age standard scores.

"Frank

Excellent replies from all. I'd like to add the legal aspect.

IDEA

300.541 Criteria for determining the existence of a specific learning disability.

(a) A team may determine that a child has a specific learning disability if:

(1) The child does not achieve commensurate with his or her age and ability levels in one or more of the areas listed in paragraph (a)(2) of this section, if provided with learning experiences appropriate for the child's age and ability levels; and

(2) The team finds that a child has a severe discrepancy between achievement and intellectual ability in one or more of the following areas:

We all know this section. Notice that the criteria is "does not achieve commensurate with his age and ability" not "his grade and ability."
I'd also like to add the point of looking closely at the

(a) A team may determine that a child has a specific learning disability if* phrase. "The team MAY determine...."

It does not require you to have a severe discrepancy. Discrepancy is inclusionary but not exclusionary.

Ron Dumont

We work in a legal environment. My state requires that age based scores be used. That's based on the federal regulations, which say:

Sec. 300.541 Criteria for determining the existence of a specific learning disability. (a) A team may determine that a child has a specific learning disability if-- (1) The child does not achieve commensurate with his or her age and ability levels

Age. Not grade. You could of course invoke Section 300.354* and argue that a severe discrepancy based on age was the result of your school having failed to instruct him. In its guidebook, "Taking Responsibility for Ending Social Promotion" (issued July 1, 1999), the U.S.D.O.E suggests that retention is never an appropriate or sound educational intervention. So you could say, "Sorry. Your kid doesn't qualify because we shafted him. Since we effectively ruined your child's chances of ever doing well in life by retaining him, he doesn't qualify for any help." If you try that argument, please let me know how it turns out.

*300.541 (b) A child may not be determined to be eligible under this part if-- (1) The determinant factor for that eligibility determination is (i) Lack of instruction in reading or math.

Guy M. McBride

While age norms should always be used to consider eligibility, there is a place for grade norms. I am not talking about age and grade equivalent scores, as some have, but about the standard scores and percentiles compared with others at the same age and the same grade level. For a retained child, I

will run grade norms on math, which is very much tied to teaching certain skills at certain grade levels. Once a child is a reader, reading and writing skills are tied less to instruction than to accessibility, so age norms are almost always the most appropriate. However, for young children in first and second grades (our cut-off is Dec. 2, as opposed to Sept. 1, so many of these children are young by most state standards) at the SST level I will run both grade and age norms to show teachers that even though these children seem "behind" compared to classmates, they are indeed acquiring skills at a "normal" rate for their age.

Rhonda J.

Actually, as much as I have personally railed against the use of grade equivalents, they also have a place. Some courts have accepted g.e.'s as valid indicators of a child's progress. When a school system is arguing in a tuition reimbursement case that it had provided FAPE to a child, pre-post grade equivalents showing, on average, a year's progress can be very persuasive. (Houston Independent School District v. Caius R. et. al., 5th Circuit, January 2000.)
I'm not as comfortable over using grade norms in an attempt to convince teachers that children with serious classroom problems are really normal. Telling a teacher a kid didn't meet state standards is one thing; implying s/he was behaving like Chicken Little when there's nothing wrong with the child wouldn't enhance teacher-psychologist rapport in my district.

I find that grade standard scores are more misleading than age standard scores. When using grade standard scores and the different curriculums from school to school system, one can not say that Johnny is achieving at the 2.3 grade level in reading is the same from school to school. It depends on the curriculum. However the standard age scores looks at where you should be academically based on your age. Therefore you can alleviate if they have been retained or not. I think the standard age score alleviates a lot of possible contributing factors. For example all children do not begin school at the same age. Just because a child is seven years old does not mean that they will be in the second grade. They may be a third grader who started school early.

That's just my $.02 worth.

I have not been paying close attention to this thread, but just wanted to add a simple FYI.

I have had considerable experience in calculating age and grade norms for major tests. I was the person who calculated all the age/grade norms for the WJR and am doing the same for the WJIII. When calculating norms, we applied psychometricians typically plot median scores for ordered age or grade groups and then smooth a curve through the data points. The curve smoothing is necessary because of the sampling error present in the small subssample statistics (each data point is typically based only on 50-100 subjects).

For what it is worth, the grade norm curve fitting/smoothing is always MUCH easier. The sample statistics show much less "bounce" (sampling error) and the curve fitting is like a hot knife going through a stick of butter. On the other hand, age norm plots always show much more sampling error between the different points and the curve fitting is often more difficult. In other words, the grade norm fitted solutions are more precise and have less error (the standard error's are smaller).

Just thought some folks may find this interesting.

Kevin McGrew

Just to clarify, Guy, I don't use this technique for children with serious classroom problems at the Student Study Teams. What usually occurs is the teacher brings up a child who is well-adjusted and making good academic progress but is "behind" most of the other children. This information, when combined with other supporting ecological information, is intended to provide a context for the child's performance as developmental rather than pathological and to identify appropriate areas of support and intervention given this situation.

In California, our teachers are so hyper about the type of emphasis placed on test results by teacher that some are referring everyone who is not at or above grade level, being as the state goal is to have all students above the 50th percentile (for grade) on the statewide assessments. Those politicians are SO smart, to be sure. Add to that the fact that principals and teachers can be removed if their school has not accomplished that goal in three years and, well, you can see that California has its
AGE vs GRADE Based scores

OWN special education system.

Rhonda

Okay, see there were these two carpenters eating their lunches while sitting under a china berry tree. One said to the other, "I don't believe in saws; I think hammers are always better." To which the other one said, "Please pass the mustard."

I probably should just stop there, but some people will not see the relevance to the "Saw vs. Hammer" debate. See it is just real hard to cut boards with age norms and hard to hammer nails with grade norms. So, next time you want to compare a child to his grade-mates as opposed to his age-mates, please remember to use a hammer.

Hubert

The only change OSERS made in the LD definition in 1997 were to address Congress's intent to expand the rights of parents to participate in the process. It recognized that there is a lot of debate over the LD definition, but its only response (both in the preface in Appendix B) was to make a commitment to study the issue over the next few years and make a recommendation.

There must be a severe discrepancy under the law for a child to qualify for services. OSERS has left that to the states to operationally define since 1977. Ron is quite right that the Federal Regulations do not and have not ever operationalized "severe discrepancy" which is why you have some states (like New York, for example, with its non defined 50% discrepancy rule) that have virtually no definition at all, some states with regression formulas, and other states with fixed cut offs.

In 1996, Wisconsin, after having applied its rules without criticism for eighteen years, came under the scrutiny of OSERS. Inasmuch as OSERS was sufficiently concerned that it withheld $52 million in state aid, and as there have been no substantive changes in its definition since, their concerns are worth reviewing in attempting to make some sense out of the legal implications of the definition.

Wisconsin's response can be accessed at: http://www.dpi.state.wi.us/dpi/dlsea/een/4-1-96ld.html

Wisconsin's interpretation, unchallenged by OSERS, is summarized in the following paragraph from their letter to Thomas Hehir, then OSERS' director:

"M-teams in Wisconsin generally recognize that the criteria contained in both the federal regulations and the state rules must guide an evaluation, but do not direct an M-team to make a finding of LD eligibility or ineligibility for a particular child. Both the state and federal criteria are permissive in that they require evaluation teams to consider certain eligibility criteria, but they do not require an evaluation team to reach a conclusion solely because the child meets or fails to meet those stated criteria. The rules require evaluation teams to consider the criteria and the performance of the child against those criteria, but they also require the evaluation teams to use professional judgment in making individual eligibility determinations."

Wisconsin's rules were idiosyncratic in a number of respects, but their contention was that MD teams misconstrued them when they were applied in a manner inconsistent with the federal definition. In some cases, children with IQs below 90 were being denied services; the federal definition only excludes children who are mentally retarded. (Which, by the way, does not exclude every child with 68 IQ from being considered LD, if their adaptive behavior scores are not significantly below average.) Wisconsin had added several areas of potential disability to the federal definition, but it required MDTs to identify two areas in order to qualify the
child; in cases where a child had a discrepancy in an area not covered by the federal definition, this was not problematic, but in a few cases some otherwise very eligible kids were being excluded. Lastly, some children with severe discrepancies were being automatically identified by some teams even though there was no finding that they needed special education in order to receive FAPE. OSERS position was that this was a flagrant violation of IDEA, and Wisconsin agreed.

It is probably not possible to list ALL the factors that a team might consider in reaching a decision, but court decisions and OSERS letters have given us some suggestions.

1. Neither a low IQ score (see above) nor a high IQ score may be used to exclude a child from consideration as LD.

http://www.ldanatl.org/bulletins/AC_1_96.html

2. The Department of Education, in its letter to LDA of North Carolina, wrote that it is "generally" appropriate for the multidisciplinary team to include in its written report (to determine eligibility) information regarding "outside or extra" instructional help or support which "may indicate the child's current educational achievements reflects the service augmentation, not what the child's achievement would be without such help." Within context, the Learning Disabilities Association interpreted that to mean a child need not have failing grades, if he or she is only passing a result of special service or support such as tutoring twice a week or a parent who spends three to five hours with the child on homework each evening. (Ref: see above.)

3. Matthew Effect. If there is prior evidence of higher IQ, and present testing shows a decline that results in the child being ineligible, the team may consider whether the disability may have resulted in significantly different learning experiences which have negatively impacted the scores. While this argument has been advanced in hearings and sometimes favorably regarded (Brody v. Dare, for example, on the wrightslaw.com website), I have not seen it argued at a circuit level. Also, while I've seen research suggesting Matthew (or Mark) effect is a real consideration, I do not have data quantifying it--that is, I have no clue as to whether one would normally expect a three point drop or a twenty point drop as a consequence. I am therefore not a strong proponent for applying this on a regular basis to every child whose score dropped (why re-test if we're going to do that?), but I do recognize it as a factor to be considered.

4. North Carolina says "When there are verbal/performance IQ discrepancies of at least 20 points on the Wechsler Scale, the higher scale IQ may be used to determine the achievement-ability discrepancy providing there is evidence that the higher score accurately reflects the student's intellectual functioning. Because of the importance of the intellectual assessment to the identification process, group intelligence tests, unjustified prorated scores or extrapolated scores and abbreviated forms shall not be used." Other state definitions may offer similar clauses.

5. North Carolina also says, "If the multidisciplinary team determines that the assessment measures did not accurately reflect the discrepancy between achievement and ability, the team shall state in writing the assessment procedures used, the assessment results, the criteria applied to judge the importance of any difference between expected and current achievement, and whether a substantial discrepancy is present that is not correctible without the provision of special education. [1999 Procedures, .1505, C. (8) (c) (iii)]" With respect to that underlined phrase, I see that as saying, "If you don't like our definition, you may use your own, as long as you write down what it is." I suspect other states, however, have similar elastic clauses. It could be especially handy when reviewing a child who meets another state standard and who would otherwise be qualified for sped under Section 504. And while it is frustrating to have your state say, in effect, "Gee, guys, if you don't like our definition, make up your own," I think that what they've written is actually consistent with and a reflection of the ambiguity in the federal regulations.

Guy M. McBride
know” or “where a child should be” based on either age or grade. What we have are learning curves with a broad range of distribution along the curve for any age or grade. Furthermore, most norm referenced tests are just that, referenced to the norm, not the criterion - curricular or developmental. Standardized tests best measure how well a student performs on that test compared to other children that age or grade across the nation (or otherwise, if state or local norms are available). Use of age or grade based norms identifies the reference group, and the child’s relative performance to that group. As we depart from that, we enter the realm of speculation, in which we try to apply research, theories, and experience to interpret the meaning of these scores.

Just a couple more cents, as Mr. Greenspan tries to hold back inflation.

Jonas Taub

I fine this interpretation of IDEA problematic. Are you say that a student whose achievement is commensurate his or her age and ability is disabled? My read of this section may be faulty but it seems to me that the intent was for lack of a discrepancy to be at least potentially exclusionary. I thought the “may” language was intended to allow teams flexibility in that just having a discrepancy does not necessarily indicate SLD eligibility. Help me to understand your interpretation.

Dick

I read IDEA to say that the determination of an educational handicapping condition is a team decision, made with adequate and accurate information. It is not, as my friend John Willis points out, ”An exercise in arithmetic...but if you use arithmetic do it right.” If a child has a severe discrepancy (and there is I believe no definition of severe or mandate that it be some numerical) then, if all else is in place (pre-referral interventions, a disorder in basic psych. process, a need for an individualized program of specialized instruction) then the child MAY be considered as a child with a handicapping condition. What I also believe (and I bet a whole bunch of attorneys do to) is that a child who does not live up to some numerical “severe discrepancy” should not be automatically excluded from being classified as a child with a handicapping condition. I think it ridiculous (but real) that a district near me has set as a criteria for severe discrepancy the need for a 24 point difference between IQ and achievement. Just yesterday I stupidly asked ”What about the child with a 23 point difference.” Answer: ”Sorry...no classification.” States, districts, and teams have the right to determine for themselves what a severe discrepancy is. Just because you dont find one doesn't mean the child cant qualify.

I am not in anyway advocating an opening of the floodgates. (We may already have that with EBD - Every Body’s Disabled or was that Emotional/Behavioral Disturbed?) I simply believe that many children with true learning problems and true learning disabilities are being unfairly, and I claim illegally, excluded from services because of the over reliance on the severe discrepancy statements. I repeat that the law says MAY determine, not must determine. May does not mean must. There are other places in the revised IDEA where the schools are allowed to do things if they choose. The language there is also MAY, not must. Look at section 20 USC 1413 (g) School based improvement plan:

"Each LEA may in accordance with paragraph (2), use funds made available...."

This doesn't demand that the LEA do it, only that it has the option to do it.

Sure there should be some "discrepancy" that triggers concern, but lack of a numerical discrepancy may be the result of inadequate assessment, improper choice of
AGE vs GRADE Based scores

tests/tasks, misunderstanding of statistical aspects (correlation, regressions), etc., Can't a discrepancy be that the child is in 3rd grade, has been given wonderful access to good teaching (avoiding the common disorder of dyspedagogia), has had interventions that seemed possible to solve the problems, and yet the child is still, based on a curriculum-based assessment, well behind where he/she should be. I say the team has the right to determine that all things being considered, this child has a severe discrepancy and is eligible for classification.

I think Shakespeare had it right when he wrote "Much Ado About Nothing." If Joseph Heller were alive today he could add Severe Discrepancy Discrepancies to his wonderful Catch-22. I think that a district that recognizes a problem in a child and refuses to classify because of the severe discrepancy clause is going to run the risk of alternative interpretations. I say, if the discrepancy is there you have an added indicator that the child is in trouble, but when the numbers dont fit, you still can classify. We would not use the clause as the sole reason to include a child, we certainly should not use it not as the sole reason to exclude.

Hope that makes sense.

Ron

Aged based scores are typically lower for those whose age is greater than average compared to grade based scores. If the average student from a given area is older than the average student in a given grade in the norm group (are we confused yet?), then the age norms are more likely to be lower than the grade norms.

In other words, Dave may be correct in his assumption that age norms are lower than grade norms if students in his district are generally older than the norm group to whom they are being compared. In a bilingual population, that may very well be true, due to the greater likelihood of retention and late enrollment. In other areas, the converse may be true.

Ron Anderson

>Go to http://www.margaretkay.com/grade11.htm

> This is Margaret Kay's site and you will find a short article called "Problems Associated with the Use of Grade Equivalence Scores. In addition to the excellent suggestions you have already received, you might want to look at the Assessment/Evaluation section at School Psychology Resources Online. In particular, look at Glossary of Measurement Terms and Some Things Parents Should Know About Testing. The latter has a simple explanation of grade equivalents and other testing concepts. two possible resources:

>http://alpha.fdu.edu/psychology/not_all_test_scores_are_alike.htm

>and

>http://alpha.fdu.edu/psychology/oat_cereal.htm
Beth,

I had the same goal this past year in my district only I didn't do it by having a large workshop. We have a language arts person who holds grade level meetings with teachers and these meetings are much more informal and relaxed. The language arts person gave me an hour at each of the grade level meetings and I put it into my schedule.

I gave a short presentation explaining the different types of test scores, what they are, what they mean, and how certain ones influence the decision making process (contribute to Type II error). Then, I used scores from the teacher's classroom reading tests and children's cumulative files from their school (without names) as examples, and we did a short exercise where we found and interpreted real scores that the teachers use every day. That made it relevant to them.

I was surprised how little background and training teachers actually had in interpreting test results. Even reading teachers who give standardized tests all the time. They were all very interested in learning, very appreciative for the chance to learn more about scores, and are using their knowledge of scores form these sessions in our Child Study Meetings and PPT's. It was a rewarding experience for everyone.

Good Luck,

Sheldon
Problems with Reading Comprehension Scores??

Among the various influences dictating the stratum of reading comprehension difficulty, all of which must be fully taken into account for a comprehensive assessment, are sentence length, word length (measured in syllables), and the grammatical or syntactic complexity and structure of the sentences.

41 words, 41 words/sentence, 5.8 letters/word: Flesch-Kincaid grade 12.0

Many things make reading harder or easier. We must look at all of them. One is how long the sentences are. Long sentences are harder to read. Another is how long the words are. Long words have more syllables. They are hard to read. A third thing is the
grammar. Complicated grammar is hard to read.

56 words, 6.2 words/sentence, 4.1 letters/word: Flesch-Kincaid grade 2.0

Good job, John! You have done it again. How do you come up with these ideas?

16 words, 5.3 words/sentence, 3.5 letters/word: Flesch-Kincaid grade 0.0

Excellent job, John! You have done it again. How do you come up with these ideas?

16 words, 5.3 words/sentence, 3.8 letters/word: Flesch-Kincaid grade 1.2

Excellent job, John!

3 words, 3.0 words/sentence, 5.3 letters/word: Flesch-Kincaid grade 5.2
"Specific learning disability" means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations.

To aid schools and school psychologists in the understanding of what a learning disability is, the federal regulations pertaining to serving the learning disabled child (Volume 42, No. 250, of the Federal Register, Thursday, December 29, 1977) provide the following commentary:

"Those with specific learning disabilities may demonstrate their handicap through a variety of symptoms such as hyperactivity, distractibility, attention problems, concept association problems, etc. The end result of the effect of the symptoms is a severe discrepancy between achievement and ability. If there is no severe discrepancy between how much should have been learned and what has been learned, there would no be a disability in learning. However other handicapping and sociological conditions may result in a discrepancy between ability and achievement. There are those for whom these (other) conditions are the primary factors affecting achievement. In such cases, the severe discrepancy may be the primary result of these factors and not of a severe learning problem. For the purpose of these regulations, when a severe discrepancy between ability and achievement exists which cannot be explained by the presence of other known factors that lead to such a discrepancy, the cause is believed to be a specific learning disability." (bold, italics added)

It would seem, based on this commentary, that you need not discover what the "disorder in one or more of the basic psychological process" is. What needs to be done is simply to prove that the ability/achievement discrepancy is not caused by something other than a learning disability! Some school districts have even written into their regulation/guideline a statement such as:

"A student with a 50 percent or greater discrepancy between intellectual potential and achievement must be deemed learning disabled unless the discrepancy can be attributed to some other handicapping or socioeconomic condition." (bold added)

Without belaboring the point too much, if you don't identify what the underlying process is that is causing the child to have difficulty, how can you help the child. The problem for the school psychologist is not in discovering that there is a problem &mdash;(the teacher has probably beaten down your door to tell you that already)&the problem is in uncovering clearly what the breakdown in process is so that some form of "remediation/intervention" can be tried.

The basic disorders that can be diagnosed in the identification of specific learning disabilities are almost endless. The following list is only
Basic Disorders

Illustrative. It should not be taken in any way as limiting the disorders that Teams can and should consider when attempting to determine the causes of school failure. The list is provided to give some idea of the scope of the possible disorders. The categories are somewhat arbitrary and are not based on any particular theory of disabilities. The order is random. Many of the disorders could be placed in two or more of the categories. The omission of any disorder from this brief list does not imply any lack of legitimacy, nor does the inclusion of a disorder imply that it carries with it some greater level of validity or importance. The list is purely illustrative.

**Auditory Perception**

- mishearing sounds
- mishearing meaningful sound units (phonemes)
- mishearing words
- misperceiving the order of sounds
- difficulty remembering the order of sounds or of words
- confusing similar-sounding sounds or words
- difficulty blending sounds into words
- weakness in auditory figure-ground perception
- difficulty hearing separate sounds in words

**Visual Perception**

- confusing similar shapes
- failing to see differences between shapes
- difficulty seeing how parts fit together to make a whole
- difficulty perceiving direction, confusing up/down, left/right
- difficulty organizing the parts of a larger configuration
- difficulty perceiving the parts when seeing the whole
- difficulty perceiving the whole when seeing the parts
- visual figure-ground confusion
- weak visual tracking
- difficulty with visual perception possibly caused by amounts of glare or contrast

**Language**

- difficulty understanding and learning the meaning of words
- difficulty understanding the meaning of sentences
- difficulty retrieving precise words quickly when needed
- difficulty understanding metaphor, irony, or figurative speech
- difficulty understanding or using prepositions
- confusion of grammatical function
Basic Disorders

- confusion of temporal sequence in language
- difficulty expressing ideas in coherent sentences
- difficulty understanding or using grammatical relationships

Motor Coordination

- deficient gross motor coordination, e.g., hopping, running
- deficient balance (static and/or moving)
- persistence of primitive reflexes, interfering with coordination
- deficient fine motor coordination
- deficient eye-hand-foot coordination (in any combination)

Sensory-Motor

- difficulties with sensory input and its integration with reflexive and purposeful motor responses
- difficulty tolerating touch or certain textures
- difficulty integrating two movements at once

Attention

- difficulty focusing attention at will
- difficulty sustaining the focus of attention
- difficulty ignoring distractions
- difficulty attending to the most important stimulus
- difficulty listening
- difficulty seeing details
- difficulty concentrating on two things at once
- difficulty recalling information when it is needed
- difficulty remembering to do things at the right moment
- difficulty making learned responses or skills automatic
- difficulty performing tasks when the situation changes

Memory

- weak memory (short-term, long-term, auditory, visual)
- weak memory for sequences
- weak spatial memory for locations
- rote memory weaker than memory for meaningful material
- difficulty associating written symbols with sounds or words
Academic skills

- primary disorder of reading, arithmetic, handwriting, spelling, or written expression not attributable to other disorders

Cognitive

- disorders of deductive reasoning
- disorders of inductive reasoning
- impulsive thinking
- difficulty forming and using categories
- difficulty generalizing
- slow mental processing (not the same as "slow learning")
- inability to think under stress
- verbal thinking abilities much superior to nonverbal
- nonverbal thinking abilities much superior to verbal
- sequential thinking much superior to simultaneous or holistic
- holistic or simultaneous thinking much superior to sequential
- "fluid" thinking ability much superior to "crystallized" knowledge
- "crystallized" knowledge much superior to "fluid" thinking ability

These are examples of the array of possible disorders that might be primary causes of severe deficits in academic skills. There can be, and often are, two or more such disorders affecting a learning disabled student.

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Common Evaluation Errors

The following is a list of common mistakes that evaluators make. It is not in anyway exhaustive and is in no particular order.

1. Looking at a child's performance in isolation, without looking at performance in relationship to schools.

2. Obtaining test scores very different from those of other evaluators and not noting the differences or hypothesizing about the reasons why.

3. Using outdated instruments or norms, without noting this and explaining why.

4. Using the wrong level instrument, or a test designed for a child of a different age or grade, without noting this and justifying why.

5. Failing to compare achievement test results with expectations based on IQ.

6. Failing to acknowledge that isolated high scores may be meaningless anomalies.

7. Re-administering tests too frequently, without acknowledging or explaining practice effect.

8. Blindly accepting testing as valid when in fact it may be methodologically flawed.

9. Being unwilling to share your protocols or other test data with competent professionals.

10. Destroying records and protocols.

11. Using experimental tests without noting this.

12. For a child who is unusually old or young compared to his/her classmates, not indicating whether the reported test scores are based on age or grade and why.
13. Trying to compress too much testing into a single session.

14. Not writing the test date on the protocol.

15. Not completing the protocol. (Filling in only the score but not the response for example.)

16. Mis-scoring the test.

17. Relying solely on a computerized interpretation of a test.

18. Recommending services for non educational reasons.

19. Identifying only weaknesses, while ignoring strengths.

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
ACADEMIC/BEHAVIOR DEFICITS

Below are listed some behavioral manifestations of what might be "disorders" in the basic psychological processes.

1. ORGANIZATIONAL SKILLS
   
   a. Difficulty adjusting to routine of classroom and school, absent often.
   
   b. Cannot organize materials, belongings, loses items.
   
   c. Disorganized work style; erratic, inexact work, messy paper work.
   
   d. Needs continual direction, reminders, and guidance in order to complete work
   
   e. Daydreams, stares.
   
   f. Cannot shift easily from one activity to another, mind on previous activity lesson.

2. ATTENTION SKILLS
   
   a. Distractible - attention paid to everything, both essential and nonessential, fiddles with pencil, paper, desk.
   
   b. Hyperactive - seems unable to stay still; constant activity: moves legs, chair, table, walks around room.
   
   c. Impulsive - apparent inability to delay or control impulses (i.e. responds without thinking, doesn't wait for directions to be given, etc.)
   
   d. Daydreams, loses track of what is going on, is in own world, etc.

3. AUDITORY PERCEPTUAL SKILLS
   
   a. Poor recall of what has been heard in past (simple words, ideas, etc.)
   
   b. Inability to follow simple oral directions and assignments.
   
   c. Uses sight-word approach to recall words but cannot attack words phonetically
4. VISUAL PERCEPTUAL SKILLS

a. Poor recall of what has been seen in recent past (sight words, numerals, pictures, etc.)

b. Does not discriminate differences in size, shape, color, perspective.

c. Can attack words phonetically but not with sight-word approach.

d. Difficulty in cutting, pasting, drawing, coloring.

e. Difficulty reproducing letters, words, designs, many erasures.

f. Loses place frequently in copying or reading, needs extra time.

5. LANGUAGE DEVELOPMENT

a. Does not tell a comprehensible story, hesitates, repeats, stutters.

b. Speaks in incomplete sentences, makes many grammatical errors.

c. Unable to relate ideas in a logical sequence.

d. Fails to grasp simple words meanings, misunderstands words at grade level, cannot recall story plot.

e. Uses immature vocabulary.

6. MOTOR CO-ORDINATION

a. Trips and falls frequently; bumps into things, unsteady awkward gait.

b. Poor handwriting, letter irregularities, reversals.

c. Limited manual dexterity for his/her stage of development, cannot tie shoes.

d. Difficulty competing with peers in sports and games because of limited agility.

e. Prefers not to participate, wants to change the rules, cries easily.
7. ORIENTATION

a. Lacks grasp of meaning of time; always late or confused.

b. Judgment of relationships is poor for age (i.e. big-little, far-close, easy-hard).

c. Confused about directions (Left-right, north-south, up-down, etc.).

d. Unable to navigate around familiar spaces (school playground)

e. Negative reaction to change.

e. "Absent-minded". Often forgets important events.

8. PEER RELATIONSHIPS

a. Is generally rejected by other children, has no friends.

b. Always teases or bullies other children.

c. Is always being "picked on" by peers.

d. Usually disregards feelings of others.

e. Does not cooperate, share, or get along with other children.

f. Prefers to associate with younger children.

g. Denies mistakes and blames others.

9. SELF-CONCEPT

a. Appears insecure or unwilling to participate in new or difficult situations.

b. Excessively shy or withdrawn.

c. Does not accept responsibility for own behavior.

d. Never assumes leadership or initiates group/class activities.

e. Easily discouraged by demands of classroom activities; gives up easily, submissive attitude toward authority.

f. Compares self to other students in negative terms.
10. EMOTIONAL STABILITY

a. Excessive anger, destructive or aggression.

b. Mood changes quickly and drastically, excitable, impulsive.

c. Constantly disrupts classroom - i.e. tantrums, unnecessary disturbing noises; always demands to be center of attention.

d. Generally very tense and nervous, worrier, overly sensitive to criticism.

e. Typically depressed and sad, withdrawn, never smiles.

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Definitions

I have always been uncomfortable with what I consider to be our inability to define well what we do and who we do it to. I think it was Voltaire who said "When you speak to me, define your terms."

The Department of Education is currently soliciting comments from concerned educators, psychologists and others, to help them decide whether to replace the current definition of SED (Seriously Emotionally Disturbed) with something with an acronym of EBD standing for Emotionally Behaviorally Disturbed.

Supposedly, the SED definition is outdated due to advances in mental health research and treatment and unfairly stigmatizes students. Moreover, it is felt that some of the terms used in the current definition such as "social maladjustment" are so vague that teachers and mental health professionals have difficulty in determining which students are eligible for services.

According to advocates, students with social or behavioral problems that affect educational performance and social interaction should be eligible for special education services. Under the new definition juvenile delinquency, substance abuse or other behavioral problems will not automatically exclude children from receiving special education services.

What does EBD stand for? I propose that if accepted the new definition will stand for Every Body is Disabled.

In 1990 I conducted a survey in New Hampshire to investigate how children were being identified as being eligible for services under the classification of SED (Seriously emotionally disturbed). At that time, based upon that survey, I proposed that no student be identified as SED in New Hampshire (Paper presented at the NHASP conference 1990). It was my belief then as now that students should not be coded, but that they should be provided appropriate services without a code. That belief was founded on the premise that we as psychologist were not uniform in our diagnosis and that the state and federal laws offered us little if any guidelines upon which to make such determinations. The government has already given us two new classification label (autism and ADD). Now we are offered a new, revised version of the old SED definition.

With the old definition of SED we only had to determine a few things (none of which we did very well):

1. Did the child have a 'condition' of emotional disturbance? What is a condition? Need it be defined by the DSM-III-R or DSM-IV or some other criteria? If not by DSM-III-R or DSM-IV then by what criteria? Did any differentiating condition need be described?
2. Once a condition was determined to be present we were to decide if it was "to a marked degree"? What describes a "marked degree"? Did we need standardized measures for this determination or just our clinical judgment?

3. If the condition existed was it "for a long period of time"? Once again, what did this mean?

4. Finally, we were asked to determine if the condition had "Adversely" affected the education of the child.

Now we are presented with a new definition. Is this one any more defining than the previous. Who won't not be found eligible under the new terms?

In the proposed new definition, "emotional or behavioral disorder" is defined as follows:

The term emotional or behavioral disorder means a disability characterized by behavioral or emotional responses in school programs so different from appropriate age, cultural, or ethnic norms that they adversely affect educational performance. Educational performance includes academic, social vocational or personal skills. Such a disability --

- is more than a temporary, expected response to stressful events in the environment;
- is consistently exhibited in two different settings, at least one of which is school-related; and
- is unresponsive to direct intervention applied in general education, or the child's condition is such that general education interventions would be insufficient.

Emotional or behavioral disorders can co-exist with other disabilities.

This category may include children or youth with schizophrenic disorders, affective disorders, anxiety disorders or other sustained disorders of conduct or adjustment when they adversely affect educational performance in accordance with section "B.I."

Under the new definition juvenile delinquency, substance abuse or other behavioral problems will not automatically exclude children from receiving special education services.

Eligibility for special education services is and should be based on the effect of the disability on a student's educational performance. However, the new definition could allow many students with various mental health or simply behavioral problems to qualify for special education services. Will the result be to diminish the resources schools have available for serving students most in need of special services without providing EBD students with the services and treatment most appropriate for their needs? Maybe the federal government could help these students most effectively by increasing the funding for community mental health programs. Is it appropriate to make schools the government entity most responsible for dealing with students' mental health problems that do not affect their academic performance?
Because the new definition does not exclude students who are "socially maladjusted" and includes students who are "conduct disordered," will schools find it more difficult to discipline students effectively? Some of the students who would qualify for services under the new definition can act very aggressively or be chronic rule-breakers. Will this new definition tie the hands of school administrators.

The new definition requires that cultural or ethnic norms be considered in classifying students. How would school staff determine what such "appropriate" norms are? Who would make such a determination?

If we don't identify these children, what are we to do? I would like to make a radical proposal to education. It will greatly increase the identification of children eligible for special education services, but in the long run will be more honest and appropriate for all students.

I propose that we identify all students passing through the public school doors as eligible for "special education" services. All students! No exceptions! Why all students? Because everyone deserves to be granted appropriate education services. I often sit in team meetings listening to the "special education" recommendations being made and added to an IEP. I sometimes protest that the modifications being made are not "special" and in fact are simply "good teaching" techniques. How often are these modifications "special." Not often! So, again, I propose we identify them all. Let's turn the special ed. process on its ear. Let's only test those who want to be considered 'normal.'

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Guide to Identification of Learning Disabilities

1998 New York State Edition

John O. Willis, Ed.D.           Ron Dumont, Ed.D.
Rivier College               Fairleigh Dickinson University

TALKING POINTS

1. SLD means a disorder in one or more of the basic psychological processes involved in understanding or in using language . . . You cannot identify a SLD without specifying a disorder and showing how it impairs school work.

2. Severity of SLD is not measured by the severity of the weakness in basic process(es). It is measured by the severity of the impact on academic achievement. Real-life academic achievement is often more important than achievement test scores. Careful, diagnostic assessment of achievement is often the core of the evaluation.

3. The "exclusions" are not as important as most people think, as long as the disorder in basic process(es) and impact on academic achievement have been documented thoroughly and properly.

4. Students may have SLD as a disability secondary to another disability, even intellectual handicap.

5. Global intelligence measured by total IQ scores (GCA, MPC, BCA, etc.) is usually not a helpful construct for understanding the cognitive functioning of students with specific learning disabilities. Part 200 Regulations do not define "intellectual ability" even as "intelligence," much less as an "intelligence test score." Cognitive ability factors are usually the most helpful level of analysis.

7. The factor structure of a test for persons with certain, specific disabilities may be very different from the factor structure for the norming sample.

8. It is almost always better to adopt an appropriate test than to adapt an inappropriate one for a student with a severe disability.

9. Parents, teachers, and the students themselves make important contributions to the evaluation, and they must be included in the process. Examiners should elicit genuine referral questions, not just "issues" to be answered by the evaluation. Interviews and questionnaires are essential parts of a complete evaluation.

10. To ensure the valuable contributions of parents, teachers, and students, evaluation results must be as clear as possible. Jargon and statistics must be defined very clearly.

11. The total evaluation must be integrated, which is not achieved with a staple.

12. SLD identification is a professional judgment by a team, not an exercise in arithmetic.

13. However, any arithmetic involved should be accurate.

14. Tests and scores are not interchangeable. A student's age- and grade-equivalent scores (which are horrible statistics) will not come out in the same rank order as the student's standard scores. Age-based and grade-based norms differ, and often both must be reported. Discontinuities between fall, winter, and spring norms can be dramatic. The same performance yields very different scores on different tests. The same grade equivalent yields different standard scores on different tests.

15. Reading tests use an extraordinary number of means of assessing reading. They are not interchangeable. Often, several reading tests must be used for a complete picture. Reading fluency and study skills are important.

16. Math tests almost always require limits-testing. Correct test scores, including "math-fact" errors and misreading of operations signs, must not be confused with the reality of the student's math skills.

17. Writing must be assessed carefully. The best formal written expression tests have many flaws. Writing samples may be needed in addition to tests.

18. Achievement testing should include detailed descriptions of actual skills, gaps, and weaknesses. It is useless simply to print a table of test scores and describe the scores in words.

19. Tests that combine separate skills in single scores (e.g., reading decoding and reading comprehension or math calculation and math applications) are as useful as a second handle on a snow shovel.

20. Discrepancy formulae are statistical nightmares. The Part 200 Regulations do not state that a "50% discrepancy" (whatever that may be)
is the only way to identify a SLD. Legal precedent indicates that the absence of a "50% discrepancy" cannot be used to exclude children from SLD classification.

21. Norms are much more important than most people think. Norms are worse than most people think. Validity and reliability matter. Validity for specific purposes and reliability over realistic spans of time are rarely documented.

22. Diagnoses are not political or economic decisions.

23. Relative and transient weaknesses must be taken seriously.

24. Examiners must use the best instruments available. Inadequate tests should be used only when they are absolutely necessary and the best existing for the purpose.

25. All disabilities, including SLD, can be seen as mismatches between learning style and instruction. Changes in circumstances can alter the need for special education.

26. Evaluation processes that always or never recommend highly restrictive placements are equally suspect.

27. Fads in diagnosis and treatment must be avoided.

28. Evaluations must be a careful, thoughtful, thorough process, whether initial or re-evaluations

29. Concrete recommendations, individually planned for the student, are an important goal of the evaluation. Stock, boilerplate recommendations are not much help. A useful evaluation with recommendations does not cost much more than a useless one without recommendations.


31. Evaluations should be individual and humanistic, consider multiple intelligences, reflect reality beyond test scores, and accept the possibility of improvement in areas of weakness as well as potential stability of individual patterns of strengths and weaknesses.

32. There is often a huge gap between the science of SLD identification and the social policy involved with the identification of SLD.

33. Best practice and educational law are often in conflict.

34. There is often a distinct difference between an evaluation for classification and an evaluation for diagnosis of educational difficulties.
Double Standards in Testing or Do We Need a Gun Control Law for Testers

Cisco (Ron Dumont Ed.D., NCSP) and Eggbert (John Willis, Ed.D.)

Eggbert and I want a Brady Bill for testers. In the real world, we require people who want to use a gun to get a permit. We do it because we recognize that guns are considered dangerous weapons in the hands of those ill prepared to use them. We suggest that tests can be the same: they are dangerous weapons in the hands of the unqualified and we need stricter laws to protect us from unfortunate hunting accidents. Some will argue that there are laws in place to protect us and others: but are they doing the job they should? Minimum protection is what the law guarantees. If making sure that tests meet standards for reliability and validity is a minimum standard, we must follow it. We believe that most school psychologists do follow this minimum standard practice, but we believe there exists a double standard. Evaluators outside the school setting are often not held accountable for the type of evaluation they do. Batteries of tests are given in non standardized fashion with little or no mention of it made in the report. We have watched the faces of evaluators turn red with anger when we have dared to question them about the tests, test procedure, and the results. God forbid a question about reliability and validity should ever come up. [If a test can't predict itself (reliability), can we expect it to predict anything else (learning disabilities)]?

Why is it that school districts are often held to a standard of testing that others are not expected to maintain. The New York state regulations (not to mention the Federal guidelines) clearly state the requirements for the testing we do. (Part 200 regulations: "School districts shall ensure that:" The law does not quibble. It states that we must ensure that minimum protections are afforded all students. Would parents, parent advocates, and attorneys, let us evaluate a child with nonstandardized tests, in a nonstandardized way, with extremely small norming samples, and decide based upon them that the child was not handicapped. Is it only when these test procedures confirm what someone else wants that "we can get away with it"? (Lest anyone get the wrong idea, we don't advocate that we try to "get away with it." Eggbert and I want us never to do it in the first place, and to hold those that do accountable for not following the law.)

As part of our position as an educational consultant to school districts and to lawyers involved in educational due process cases, we have reviewed quite a few reports and evaluations done by others. As we go through these reports they range from excellent to highly questionable. For example, we reviewed a test report so large it started with a table of contents. It contained over 45 pages of test explanations. (The child had been administered 64 separate tests). At the end of the report came the four recommendations, only one of which concerned the child: "Further testing to clarify the issues raised by this evaluation is needed" !!! To steal a phrase directly from Eggbert "the integration of this 45 page report was done with the staple." Here the issue may have been too much testing with too little integration.

In another case, we had not been allowed to evaluate the extensive report before the team meeting. When the report was finally handed out to the team, we simply held it to our forehead and did a bad "Karnac the Magician" imitation. "And the answer is..." Without ever reading it, we were both able to repeated verbatim the recommendations made (something akin to the performance of Tweedledee and Tweedledum).
parents were amazed, but still believed everything presented in the report. Maybe here the case was too little interpretation.

Finally, a child, tested by her school district, was found not to have any serious learning disabilities and therefore was not identified as handicapped. (She was though, offered non-special education services to address some of the issues highlighted by her evaluation.) Her parents were encouraged by a child advocate to take her for further testing "by the experts." They could surely find the problem that the school had overlooked. When she arrive at the evaluators office, the evaluator was ready. Having reviewed all the pertinent background data, the evaluator knew exactly what the referral problem was and what areas needed to be assessed. Parent interviews had gleaned a lot of information about the child, and the testing done by the school was also reviewed. Over the next 6 hours, the child was assessed in the areas of intellect, neurodevelopment, academic achievement, and social emotional stability. Tests include the WISC-III (her second in 3 weeks), parts of the Stanford Binet 4th Ed., subtests of the DTLA-III, the TOLD, Bender, Rey-Osterreith, parts of the WRAML, the CAVLT, the WMS-R, Menyuck Sentences, Trails A and B, the Stroop, Cancellation tests, the Hooper, the Money Road Map Test, Sentence Completions, TAT, the Rorschach, Woodcock-Johnson Psychoeducational Battery (Cognitive and Achievement), The PEERAMID, and the WRAT-R. After this evaluation, the report developed by the evaluator stated that there were "significant learning disabilities", and "that these undiagnosed disabilities created an 'assault' on the child's emotional well being."

How does one respond to such a thorough evaluation by a "competent expert". First, we must accept that the expert may in fact be right and that the evaluation done by the school had missed the boat completely. When there is such divergent opinion on both sides, one way to solve the dilemma is to evaluate the evaluation. This is necessary but unfortunately leads to "dueling psychologists", and seldom ends in agreement between the parties.

There are a number of ways to review a report but this essay focuses on minimum protection. Focusing on only the third example above, I believe that minimum protection was not offered to the child and therefore the results of the evaluation are questionable from both a legal as well as psychometric points of view.

Obvious issues to raise would include: the effects of test retest. Did the expert, who knew what tests had been given by the school, repeat them after a short period of time and not account for the typical change in scores? Besides the obvious fact that any child taking this many tests in such a short period of time will be fatigued and probably overwhelmed, what is the need for such an extended battery? What is the effect on subsequent test scores of adding more and more tests to an evaluation? Does the adding of test after test increase the reliability of discovering the problem? (Cisco once challenged an opposing lawyer to let him test him for just 10 minutes. Cisco would guarantee to "find" something wrong. It's really just a matter of choosing the right test. The lawyer declined.) Is there such a thing as incremental validity and is it something that we should strive for? Since an individual's ability to synthesize and understand all the small incremental pieces of knowledge is extremely limited, what is gained by giving such a 'complete' battery? It could be that what is gained is test/testee bias, confirmatory bias, illusory correlations, etc.. How often do we receive a request to evaluate and then find nothing wrong? We are the experts; we're supposed to be able to find and explain the difficulty. Do we find difficulty or simply confirm what others told us to find? Isn't testing and evaluations supposed to be objective? Shouldn't every student who comes to us be viewed as normal, and we are out to prove it?

"Tests and other assessment procedures: (b) have been validated for the specific purpose for which they are used; "

Review the list of tests given. We would characterize some of the tests as esoteric. (That's to choose a nice word for what we really think of some of them.) What are the psychometric properties of each? The Stroop for example: Created in the 30's and renormed in the 70's by
Charles Golden. One evaluator gave this test to a 7 year old and diagnosed "diffuse bilateral brain damage." Serious stuff for a test with extremely questionable statistics. The data used to create the age based norms were lost when Dr. Golden moved from California (personal conversation 10/89). To his credit, the manual states that the norms are provisional and need to be updated. Unfortunately this hasn't happened. Does it matter? Cisco's study of the Stroop found the age corrected norms to be so far off that a normal person will receive scores placing them from .3 to .9 standard deviation units below the mean. How stable and reliable are the norms for some of the tests? Remember, the standards demand that the test be reliable and valid for the purpose for which they are given. Although the standards don't say that the responsibility for choosing tests that are reliable and valid ultimately lies with the evaluator, who else's responsibility should it be? It must also be noted that this does not mean, as some have suggested, that the evaluator creates these properties. We once sat open mouthed at a special education hearing as we listened to an evaluator explain that because s/he was an expert tester, the tests and results were therefore valid and reliable. This person seemed to think that the tests become reliable just by giving them properly. Tests don't become reliable, they either are or aren't based upon the statistics. How many people are included in the norming samples of tests like the Cancellation test, the Trail Making Test, or the Rey-Osterreith Complex Figure task? Often it will depend upon whose norms you wish to use and choosing norms can make a big difference in interpretation. A score on the Rey using the original norms will place an 8 year old well below average, yet the same score using the Whishaw-Kolb norms places the same child at dead average.

"and (c) are administered by trained personnel in accordance with the instructions provided by those who developed such tests or procedures."

Looking at the list of tests given one notes that in a number of cases only selected subtests of larger tests were given. Is this minimum protection? What happens when tests are mutilated? How can an evaluator claim reliable results when the test is given in an unstandardized way? What other nonstandardized practices can be allowed? Can we disregard starting and stopping rules and still compute standard scores? Can we raise or lower the scores of children simply by asking probing questions? As soon as we change from standardized administration to nonstandardized administration, we corrupt the resulting score. If we don't account for this change in the report, we have not met legal and ethical standards. Even if we do report the changes, can we now report anything about the reliability of the subtest?

Use of certain tests and their results with a suspected learning disabled child can raise another whole set of problems. Does a language disability or perceptual difficulty effect the way a child might respond to the projective stimulus of the Rorschach or other projective tests? For the learning disabled child will perception alter the output and thus effect the scoring? A review of the PSYCHLIT data base, running a search for Rorschach and Learning Disabilities, found a total of only 2 articles written addressing the problem. Some would suggest that we just not score the test protocols, and that we focus on the 'process.' That's fine to a degree. The process is certainly important when we evaluate children, but what if the process, no matter how difficult or revealing, results in "average" scores? Is a child disabled because the process is difficult but the results are normal?

Are evaluations done with tests that don't meet minimum protection standards harmful to children? Why should we (and everyone else) be held accountable for the testing we do? What's wrong with giving esoteric tests, or tests with no norms, or norms so abysmal that a child will be compared to 5 others to determine abnormality? Why not use a test that has no reliability or validity studies? What harm is done? Why not pick and choose subtests from larger tests, give them, and consider them valid measures of what they say they measure? Why not administer tests in a nonstandardized manner? Is the only reason to do it because the law says to?

The answer to these questions is fairly simple. Yes, it is the law and that's the reasons to be held accountable, but beyond that, it's the child.
it fair to evaluate a child using poor assessment practices or poor assessment tools and identify strengths and/or weaknesses when they don't exist? We say no. It is fair and ethical to identify strengths and weaknesses using highly valid and reliable tools. From these and only these can real educational remediations be made. Don't we harm a child by trying to remediate a weakness when it doesn't exist? Or try to use a strength as a compensation when the strength isn't there in the first place.

Eggbert and I have developed a list of questions that we feel any examiner who evaluate children (and those who might need to protect the children) should be able to ask and answer. Here are just a few:

- What is reliability? How do you know if it's high enough to be useful? Is it created by the test or the tester? Why is it important?

- What is validity? How do you determine it?

- Can a test be reliable and not valid, or valid but not reliable?

- What was the standardization group? How many and from where? Was it a random stratified group or something else? Was enough data compiled to make reasonable assumptions based upon SES and geographic locations?

- If a cognitive battery has been given, can the person who gave it define the construct it measures. (What is intelligence?)

- What is subtest specificity? Does the subtest you are interpreting have adequate specificity?

- What is factor analysis? Is the test being used factorially valid?

- Whose norms were used? How old are the norms? Into what discrete reference groups are the norms broken down?

We present two opposing proposals about the assessment of children. Here they are:

**Proposal 1: Accountability.** We would propose that we all be held accountable for the testing we do. If we evaluate with tools that don't meet minimum requirements, let's say so, or do something about it! Let's be sure the parents understand how we arrived at our decisions. Let's require the presentation of norms if asked, and even if not asked. If we can't produce reliability and validity statements about the tests we use, let's be forthcoming with that knowledge. Remember the quote "In God we trust, all others bring the data." Let's not be shy about sharing protocols, or discussing second opinions. If we test properly, why not share the protocol. If someone else interprets it differently, so be it, but the protocol itself ought to be fair game. If we make a statement about a child based upon some response or group of scores, we should be willing to share the exact response and/or scores with anyone qualified or willing to interpret them. How often do we see Rorschach interpretations with no scoring sheet attached? We are willing to give WISC-III subtest scores, why not Rorschach responses and scores? We must be held accountable for our evaluations, and they should and must meet state standards. These are minimal standards. Why can't the minimum be met? If we meet them, so too must all evaluators.
Proposal 2: NONSENSE. A way of testing that uses a simple approach. This approach does not meet federal/state standards but others are using it so why can't we! Following are the eight rules of NONSENSE:
Suspension Notes

Suspension is temporary.

Basic due process rights are required.

If suspensions are repeated and the cumulative amount of time suspended approaches 10 days, it may constitute a 'change in placement' and require complete due process proceedings.

If a series of suspensions that are each of 10 days or were in duration create a pattern of exclusions that constitutes a "significant change in placement", the requirements of 3 C.F.R. Sec. 104.35(a) also apply.

Ponder:

- length of suspension
- proximity of suspensions to one another
- total amount of time the child is excluded from school

Time out:

Parents of students in special education program brought action against school officials and school district arising out of placement in time-out rooms for in-school suspensions.

Dennis and Sally Hayes were placed in the PSA program for the 1980-81 school year. During that year both children behaved in a disruptive manner and violated school rules. Consequently, the children were required, at various times, to stay in a three-foot by five-foot room for "time-out" periods and in-school suspensions. School officials placed students in the time-out room for in-school suspension as a method of punishment, and for short "cool-down" periods "to ensure the safety of other students in the classroom from disruptive behavior." Hayes, 669 F.Supp. at 1528. Rather than sending the children home as a form of punishment, the school preferred to keep suspended students at school so that the children being disciplined would not "just roam the streets and goof off." Id. at 1527. The time-out room was located in an annex of the classroom so that supervision over the student could be maintained while the student remained in the room. The district court specifically found that "the school's use of the time-out room ensures that [the students] would not be deprived of their educational rights while they were on suspension," and that "when a student was placed in the 3' x 5' room for in-school suspension, that student was to be working on classroom material." Id. at 1528. [2] This case is illustrative of the close relationship between the use of discipline and in-class instruction in providing a child with a "free appropriate public education." We believe that "proper conduct and education
are inextricably intertwined," DOE by GONZALES v. MAHER, 793 F.2D 1470, 1491 (9th Cir.1986) and hold that the discipline of a child in the classroom, including short-term suspensions and "time-out" periods, is a matter that relates to the public education of a handicapped child and that therefore falls within the scope of the EHA.

Due process:

KAELIN v. GRUBBS

Plaintiff, Michael Kaelin ("Michael"), was a 15 year old ninth grade student at the Walton-Verona Public Schools during the 1978-79 academic year. He has been identified as a handicapped child since kindergarten. In August 1978, the Walton-Verona High School identified and evaluated Michael as a "handicapped or an exceptional child". Pursuant to an Individualized Education Program ("IEP"), he was placed in an Educable Mentally Handicapped ("EMH") classroom. On March 13, 1979, Michael defied the authority of his teacher, William C. Daniel ("Mr. Daniel"). Michael refused to complete assigned classroom work. He also destroyed a work sheet and one of Mr. Daniel's coffee cups. Moreover, in attempting to leave the classroom, Michael pushed, kicked and hit Mr. Daniel.

The next day, Michael was suspended from school. On April 17, 1979, the Walton-Verona Board of Education ("Board") held a hearing concerning Michael's behavior. The Board did not convene or consult the AARC before or during this hearing. Moreover, the Board did not address the relationship, if any, between Michael's handicap and his disruptive behavior. On April 18, the Board concluded that Michael had violated Ky.Rev.Stat. s 158.150 [FN3] and Walton-Verona Board of Education Policy 609.1. Consequently, the Board expelled Michael from school for the remainder of the 1978-79 school year. This expulsion was effective April 30, 1979.

In this case, Michael was expelled without receiving the procedural protections afforded by the Handicapped Children Act and Section 504's implementing regulations. The Board did not address the relationship between Michael's disruptive behavior and his handicap.

(2) Change in placement. Change in placement refers to those actions that cause a significant alteration in programming for a child who is currently receiving special education and related services. Such alterations may be admissions and release committee initiated or may be the result of extenuating circumstances (e.g. family moves).

(a) The following actions shall be considered significant alteration or change in placement for an exceptional child. A change from:

1. Special education and related services to regular education, including regular education with support services;

2. One categorical program to another (e.g., TMH to EMH);

3. One program plan to another (e.g. special class to resource room);

4. One instructional level to another (e.g. elementary to middle school);

5. A special school or setting to or from a regular school;

http://alpha.fdu.edu/psychology/suspension_notes.htm (2 of 5) [6/12/02 3:34:28 PM]
6. One school district to another school district;

(b) Any change in placement shall follow due process procedures to insure that exceptional children and their parents are guaranteed procedural safeguards in decisions regarding identification, evaluation, and placement, including written parental permission for change in placement.

(c) Any change in placement shall be subject to established admissions and release committee procedures and consideration of the least restrictive environment concept.

Procedural safeguards

1. written notice
2. notice include reasons and rights
3. impartial decision maker
4. appeal to state education agency

Suspension and exclusion of handicapped students:

In S-1 v. Turlington, 635 F.2d 342 (5th Cir.), cert. denied, --- U.S. ----, 102 S.Ct. 566, 70 L.Ed.2d 473 (1981), the plaintiff school children were mentally handicapped. These plaintiffs were expelled from school for alleged misconduct without receiving the procedural protections required by the Handicapped Children Act, Section 504, and its implementing regulations. The court held that "an expulsion must be accompanied by a determination as to whether the handicapped student's misconduct bears a relationship to his handicap." Id. at 346. The state and local school authorities have the burden of determining whether a student's misconduct is a manifestation of the student's handicap. Id. at 349.

The court did not hold, however, that handicapped children could never be expelled from school. Rather, "expulsion is still a proper disciplinary tool under the (Handicapped Children Act) and section 504 when proper procedures are utilized and under proper circumstances." Id. at 348.

Unlike the expulsion of a non-handicapped child, however, "the complete cessation of educational services during an expulsion period" could not be authorized for a handicapped child."

An expulsion is a change in educational placement within the meaning of the Handicapped Children Act.
Stuart v. Nappi, 443 F.Supp. 1235, is the seminal case which addresses this issue. The plaintiff, a high school student with complex learning disabilities, limited intelligence, and a history of behavioral problems, was suspended from school following her involvement in several school-wide disturbances. She contested a recommendation that she be expelled, requested a hearing and sought review of her special education program. The court enjoined the school board from conducting a hearing to expel the plaintiff. The court reasoned that the Handicapped Children Act prescribed a procedure for transferring disruptive children. See 42 C.F.R. s 121a.522. Only a professional team analogous to Kentucky's AARC could change a handicapped child's placement.

The court noted that there was a conflict between the procedures required by the Handicapped Children Act and the disciplinary procedures of the local school. The court, however, relied upon a comment contained in 45 C.F.R. s121a:

While the placement may not be changed (after a complaint proceeding has been initiated), this does not preclude a school from using its normal procedures for dealing with children who are endangering themselves or others. The court interpreted this regulation and the Handicapped Children Act as prohibiting disciplinary measures which effectively change a handicapped child's placement. The court also reasoned that the "right to an education in the least restrictive environment may be circumvented if schools are permitted to expel handicapped children." Id. at 1242. Therefore, the court concluded that the "use of expulsion proceedings as a means of changing the placement of a disruptive handicapped child contravenes the procedures of the (Handicapped Children Act)" Id. at 1243.

The court in Stuart v. Nappi, however, did not hold that handicapped children were immune from school discipline.

The court stated:

Handicapped children are neither immune from a school's disciplinary process nor are they entitled to participate in programs when their behavior impairs the education of other children in the program.

First, school authorities can take swift disciplinary measures, such as suspension, against disruptive handicapped children. Secondly, a (special education committee) can request a change in the placement of handicapped children who have demonstrated that their present placement is inappropriate by disrupting the education of other children. The Handicapped Act thereby affords schools with both short-term and long-term methods of dealing with handicapped children who are behavioral problems. Id.

In Doe v. Koger, 480 F.Supp. 225, the plaintiff, a mentally handicapped student, was expelled for disciplinary reasons. The plaintiff requested and was denied a special hearing used in the placement of handicapped students. The court held that the language of the Handicapped Children Act and the accompanying regulations indicated clearly that the Act was intended to limit a school's right to expel handicapped students. The court noted that neither 20 U.S.C. s 1415 nor any of the accompanying regulations provide for the expulsion of handicapped students. Moreover, schools were not to expel students whose handicaps caused them to be disruptive. Instead, the schools were to appropriately place these students in a more restrictive environment. Therefore, the court concluded that a handicapped child could not be expelled if his handicap caused his disruptive behavior.

The court stated:

For an appropriately placed handicapped child, expulsion is just as available as for any other child. Between a handicapped child and any
other child, the distinction is that, unlike any other disruptive child, before a disruptive handicapped child can be expelled, it must be determined whether the handicap is the cause of the child's propensity to disrupt. Doe v. Koger, 480 F.Supp. at 229.

We adopt the analysis contained in Turlington. Accordingly, we hold that an expulsion from school is a change in placement within the meaning of the Handicapped Children Act. Under the Turlington analysis, a handicapped child is not totally immunized from disciplinary action by the Handicapped Children Act, Section 504, or supporting case law which holds that an expulsion is a change in placement. First, it is well-settled that a handicapped child may be suspended temporarily without employing the procedures in 20 U.S.C. § 1415. Stuart v. Nappi, 443 F.Supp. at 1242; Sherry, 479 F.Supp. at 1337; Doe v. Koger, 480 F.Supp. at 229. Second, as long as the procedural protections of 20 U.S.C. § 1415 are followed, a handicapped child may be expelled in appropriate circumstances. Turlington, 635 F.2d at 348; H. R. v. Hornbeck, 524 F.Supp. 215, 219 (D.Md.1981); Doe v. Koger, 480 F.Supp. at 229. A handicapped child may not be expelled, however, if his disruptive behavior was a manifestation of his handicap. Turlington, 635 F.2d at 348; Doe v. Koger, 480 F.Supp. at 229. Moreover, even during the expulsion period there may not be a complete cessation of educational services. Turlington, 635 F.2d at 348. Therefore, handicapped children can generally be disciplined in the same manner as nonhandicapped children.

"It does not embrace conduct that bears only an attenuated relationship to the child's handicap. An example of such attenuated conduct would be a case where a child's physical handicap results in his loss of self esteem, and the child consciously misbehaves in order to gain the attention, or win the approval, of his peers. Although such a scenario may be common among handicapped children, it is no less common among children suffering from low self esteem for other, equally tragic reasons." Doe v. Mather, 793 F.2d, 1470

Graduation and 'change of placement:

CRONIN v. BOARD OF EDUCATION OF the EAST RAMAPO CENTRAL SCHOOL DISTRICT;

Handicapped student and student's parents brought motion for preliminary injunction to require school district to continue student's educational placement pending final judgment as to whether student should remain in his educational placement or be graduated. The District Court, Robert J. Ward, J., held that: (1) decision to graduate handicapped student was "change in educational placement" under Education of Handicapped Act that triggered all procedural protections of Act; (2) handicapped student's removal from educational program by graduation during pendency of administrative proceedings violated stay-put provision of Education of Handicapped Act; and (3) handicapped student was entitled to preliminary injunction to require board of education and Commissioner of Education to reinstate handicapped student to educational program.

Injunction granted.
POST-OTITIS AUDITORY DYSFUNCTION

John O. Willis, Ed.D., Rivier College, 2/16/98

Post-Otitis Auditory Dysfunction (POAD) is a serious and underestimated cause of multiple learning problems in school children. Despite Boucher's (1986) dramatic findings, the problems of POAD often go unrecognized and untreated.

Middle Ear Dysfunction

Episodes of middle ear fluid build-up (serous otitis media) are a common plague of early childhood. Because young children often are not aware of the blockage and because the episodes are not life threatening, many such episodes go untreated or are treated too little and too late. Fluid accumulations may remain, blocking hearing, after any painful and visible symptoms have disappeared. A new episode may begin soon after the previous one and not be noticed. One consequence of these events is that a child may be deprived intermittently and unpredictably of usable hearing during critical stages of language acquisition.

That intermittent and unpredictable hearing loss can, at its worst, interfere with acquisition of basic oral language skills, both vocabulary and grammar. More subtle effects can include deficiencies in auditory perception and development of phonemic awareness or the ability to recognize the separate sounds that make up a word, skills that are essential for the development of reading and spelling skills (e.g., Brody, 1994; Stanovich, 1994). Other essential auditory processes may be impaired, such as the ability to hear against background noise and the capacity to sustain listening attention, even when highly motivated. The central nervous system tends to abandon sensory systems that prove ineffective during critical developmental periods. For example, a strabismus (deviation of an eye off-center), causes overlapping, conflicting visual projections in the visual cortex of the brain, which eventually copes by shutting down the input from the lazy eye, resulting in amblyopia, a dimness of vision which cannot be corrected by prescription lenses. This self-destructive reaction is sometimes prevented by patching the straight eye so that the off-target eye is forced to work alone, sending a single, unambiguous projection to the brain. Similarly, intermittent hearing loss can sometimes severely impair development of listening skills. The young child, without knowing why, sometimes is able to hear and understand and sometimes is not. The child begins to learn that hearing is, at best, an unreliable process, and the development of listening skills, auditory attention, and auditory perception is impaired. Attention span and organizational skills may also develop poorly, mimicking Attention Deficit Hyperactivity Disorder (ADHD).

Effects of POAD on Academic Achievement

The impairment of communication skills has many consequences. Not only do auditory, language, pre-reading, and pre-writing skills suffer
delays, but the child also suffers confusion and embarrassment from frequent social misunderstandings. The child is likely to spend considerable time not really understanding what is going on and therefore often saying and doing the wrong thing. Public ridicule is often the lot of the child with middle ear dysfunction, sometimes resulting in extreme shyness.

If episodes of fluid build-up continue into the child's school years, the consequences can be especially severe. Confusion, misunderstanding, and embarrassment are exacerbated by difficulties with beginning academic skills. Even if the child no longer suffers from episodes of hearing loss, the problems with auditory perception, phonemic awareness, vocabulary, grammar, listening skills, hearing against background noise, and attention and organization persist, putting the child at risk for school failure and further embarrassment and confusion. Compounding the problem is the evidence that the child can hear, at least some of the time, even if the fluid build-up continues into grade school.

Even very sympathetic teachers find it difficult to believe — certainly to remember — that the child has serious auditory problems, since the child does hear adequately, under ideal listening conditions, between episodes of fluid build-up and after the episodes have finally ceased.

Boucher (1983) used pure-tone and impedance audiometry records to study all of the readiness and primary students with middle ear dysfunction in ten New Hampshire towns. The students whose impedance testing showed evidence of previous otitis, even though they had regained normal hearing, showed double the normal rates of readiness placements, disability identifications, and repeated grades.

Ways of Teaching the Student with POAD More Effectively

The most prudent and helpful approach for teachers is to assume that the student has listening difficulties at all times. Even after the episodes of fluid build-up have long since ended, the continuing problems with auditory perception, phonemic awareness, vocabulary, grammar, listening skills, hearing against background noise, and attention and organization make it wise to work with the student as if the student actually were hard of hearing and as if the student had Attention Deficit Hyperactivity Disorder.

It is helpful to insist that the student sit where the teacher can be seen and heard most easily (not always front-row-center, depending on the teacher's habitual movement in the room). Without further embarrassing the student, you can inconspicuously make frequent eye contact and call on the student when you are certain the student is ready and able to respond.

Visual teaching methods and materials can be extremely helpful. Charts, maps, graphs, diagrams, models, time lines, dioramas, illustrations, demonstrations, hands-on activities, and role-playing may be effective means of teaching the student and allowing the student to demonstrate mastery of the material.

It is essential to evaluate the student's oral and written work for indications that the student is missing some essential skill or piece of information and to remedy those deficiencies. Students with POAD almost always miss out on key facts and skills, deficits which haunt them as they try to progress through the higher grades. If the gaps are too large or too pervasive to handle with brief instruction, the student may need specific tutoring from you or from the special education department.

If the student struggles with reading, or in the higher grades, reads slowly or inefficiently, the most probable causes are, in order of probability, deficient phonemic awareness, deficient phonetic word attack skills, limited reading vocabulary, and difficulty comprehending
complex written language. You may be able to identify the problem by working individually with the student or you may wish to refer the student for a comprehensive reading assessment. The student may require remedial reading services.

Examine the student's tests and quizzes to see if the student may require additional time or a quieter environment for taking tests. That should not be an automatic response for all students, but may be helpful for some.

Despite your best efforts in class, the student is very likely to miss essential details about assignments, long-term projects, upcoming events, and missing assignments. For the student to derive the full benefit of your instruction, it is essential to check in with the student (or have the student's case manager do so) frequently to be certain the student is aware of what needs to be done. It is highly likely that student will be too shy or embarrassed to initiate the checking. It is a valuable long-term goal, especially in the higher grades, for students to become independent in monitoring their own assignments and progress. However, that will not happen overnight, and it will not happen without direct assistance from teachers by gradual steps. Simply telling the student to be more responsible or giving failing grades will not overcome the lifelong effects of POAD.

Students with POAD often need additional explanations of material that has been presented in class. It is safe to assume that if any of your students with other learning disabilities or ADHD are having difficulty with some of the material, your students with POAD will also be confused, even if they are too shy to ask, are too confused to know what to ask, or are too discouraged to try without encouragement.

It is important to be gentle with students who have POAD. They have spent a lifetime being criticized and yelled at. They need support and encouragement, and most respond very well once they come to believe and trust what may be to them a very unusual and unexpected approach.

Students with POAD present teachers with special challenges and require hard work, but it can also be very rewarding to help a struggling student achieve his or her potential in your class.

**Special Education Identification**

If the student's POAD seriously impairs academic performance and is the primary cause of that impairment, if the student is no longer suffering a current hearing loss, and if the student requires a uniquely designed program of special education, the student is eligible for identification as having a specific learning disability (NH Standards, 1996, pp. 10-11, 26-27; Willis, 1990). If the other considerations apply, but the student needs classroom modifications without a uniquely designed program of special education, it might be more appropriate to provide the student with a plan under Section 504 of the Rehabilitation Act of 1973 (PL 93-112).

**References**


http://alpha.fdu.edu/psychology/poad.htm (3 of 4) [6/12/02 3:34:36 PM]
New Hampshire standards for the education of students with disabilities (10/21/96). Concord, NH: NH Department of Education.


Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
The **Overall Achievement Test: Cumulative Evaluation Reflecting Educational Ability Levels** (sometimes referred to as simply: OAT-CEREAL) was created by Dr. John Willis to illustrate one aspect of the almost total mindlessness of grade equivalents for the reporting of educational test data. The apparent clarity and simplicity of these statistics are deceptive traps for the unwary. At first glance, grade equivalents seem to reflect, with a commendable return to basics, a student's actual functioning level in a given subject area. However, since grade equivalents are simple transformations of raw scores, a student may in fact do no work at all on the grade level reported. The child might do unusually accurate work below the reported level-or the opposite.

Take, for example, Ralph, who tends to be very careless with simple computation and has a slight tendency to confuse the numbers, 6 and 9. Consequently, on the OAT-CEREAL Math test, he failed items 1, 2, 3, 4, 5, 6, 7, and 9, and obtained a raw score of 4 for a grade equivalent of 4.5. This despite the fact that he had no difficulty whatsoever with items 8, 10, 11, 12, all far above a fourth grade curriculum.

**Math**

1. 6 + 2  
2. 9 - 6  
3. 43 - 19  
4. 26 x 4  
5. 329 ÷ 8  
6. 3571 ÷ 63  
7. 2/9 + 5/9  
8. 7/12 - 2/3  
9. 3.964 ÷ 2.91  
10. 3x + 2y = 8 Find x  
11. factor: x² + 2xy + y²  
12. …3x²dx

**Math Normative Data**

<table>
<thead>
<tr>
<th>Raw Score</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
</table>

http://alpha.fdu.edu/psychology/oat_cereal.htm (1 of 3) [6/12/02 3:34:46 PM]
Bunny, an otherwise very competent reader, still confuses the letters, t and f. This unfortunate error caused her to hilariously mispronounce eight of the twelve words on the OAT - CEREAL Oral Reading Test and obtain a grade equivalent of 4.5, although she read (and understood) without difficulty the words, cacophony and polyandry.

**Oral Reading**

1. cat
2. sink
3. crashing
4. touches
5. plastics
6. debatable
7. trough
8. magistrate
9. remonstrate
10. readministration
11. cacophony
12. polyandry

**Reading Normative Data**

<table>
<thead>
<tr>
<th>Raw Score</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Equivalent</td>
<td>1.2</td>
<td>2.3</td>
<td>3.4</td>
<td>4.5</td>
<td>5.5</td>
<td>6.5</td>
<td>7.5</td>
<td>8.5</td>
<td>9.6</td>
<td>10.7</td>
<td>11.8</td>
<td>12.9</td>
</tr>
</tbody>
</table>

The point -- obviously enough -- is that, while grade level designations of test items or groups of items may be meaningful and even helpful from time to time, especially with criterion-referenced measures, grade equivalents of raw scores not only fail to present any meaningful information but may in fact serve to convey impressions that are entirely contrary to fact. For normative purposes, it is therefore much wiser to compare student's test performance to that of the student's peers through standard scores, percentiles, stanines, normal curve equivalents, or similar ranking or deviation measures.

"Statistics don't lie, but statisticians do."

For more information, see:


As you can see, John's been at this for a while!

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
LOGICAL STEPS IN DETERMINATION OF A SPECIFIC LEARNING DISABILITY

Is there a problem with academic performance?

A. Does the student have low scores on group or individual achievement tests?
B. Is the student receiving low or failing marks in a class?
C. Is the student working much too hard or much too long to earn adequate marks?
D. Is the teacher making extraordinary adaptations or accommodations for the student?
E. Is there a specific area of performance (e.g., tests, homework, note-taking, etc.) that is notably deficient?
F. Is there another indication of insufficient academic performance?

Are there one or more disorders in basic psychological processes involved in understanding or in using language, spoken or written?

A. Can each disorder be observed or inferred from academic performance?
B. Can each disorder be documented through assessment?

Can the Team make a logical argument that each identified disorder manifests itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations?

What is the best estimate of the student’s actual intellectual ability?

A. Which scales, factors, or subtests on intelligence tests are likely to be depressed by the disorder or disorders?
B. Which intelligence test, scales, or factors would be likely to yield an estimate of actual intellectual ability uncontaminated by the disorder or disorders?
C. What is the best estimate of the student’s actual intellectual ability?

Problems may be subtle or difficult to document, but if there are no academic problems at all, there is no educational disability. [A problem with an important life function other than academic performance might trigger an identification under Section 504 of P.L. 93-112 or the Americans with Disabilities Act (ADA).]

Pay close attention to reports of problems that do not cause low marks. For example, the teacher may already be providing an informal program of special education.

This step follows next in a logical sequence, but determination of any disorder(s) may not be clear until completion of psychological, educational, speech and language, occupational therapy, physical therapy, or other evaluations. There should be multiple confirmations of any disorders.

It is not enough simply to specify deficient achievement and a disorder. There needs to be a logical, cause-and-effect relationship between the two.

See Mark 4:25. The Team must not allow a disorder to depress estimates of both intelligence and achievement and then conclude there is no discrepancy between the two. For example, verbal and nonverbal learning disabilities, respectively, will depress verbal (Gc) and visual, spatial (Gv) intelligence measures. For another example, a disorder in quantitative knowledge (Gq) would depress the WISC Arithmetic and Verbal IQ scores and DAS Sequential & Quantitative and Nonverbal (fluid) Scale scores.

Remember that achievement may be assessed by means other than
Logical steps in LD Determination

Based on those measures?

Is there a severe discrepancy between the student’s level of intellectual ability (4. C.) and the student’s achievement in:

A. oral expression;
B. listening comprehension;
C. written expression;
D. basic reading skill;
E. reading comprehension;
F. mathematics calculation; or
G. mathematics reasoning?

Does the student require special modifications of or accommodations in the educational program in order to achieve at levels commensurate with age and ability (4. C.)?

Does the student require a uniquely designed program of special instruction in order to achieve at levels commensurate with age and ability (4. C.)?

test scores (I. B. – I. F.). Maintain a bias in favor of reality. Achievement tests must be chosen thoughtfully. For example, a very brief achievement test is not a valid measure of academic performance for a student with a short attention span, and an untimed, silent reading test will not pick up problems with reading fluency. Do not obsess over formulae. Some data will not fit formulae. The Team must exercise reasoned, professional judgment.

Here is the crucial issue for identification under Section 504 or the ADA. The needed accommodations or modifications should be more than you would routinely ask of a teacher of moderate skill, experience, and dedication for any student.

This is the crucial issue for identification of an educational disability. If the student does not require a uniquely designed program of special instruction, but meets the other criteria, the identification should be under Section 504 rather than the Individuals with Disabilities Education Act.

John Willis
Rivier College
5/17/99

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
When determining predicted achievement, the Table provided take into account both the reliability of each measure used and the correlation between the ability and achievement.

The correlation between two tests is important when defining a severe discrepancy. Tests of ability and achievement do not correlate perfectly (measure the same thing). Just because someone has a low ability score, we should not assume that they would also, automatically, have an equally low score on some other, different measure. No one would assume that just because a person had a very high IQ that the person would obviously be able to excel at gymnastics, or be able to hold his or her breath longer than "normal." Just because someone is good or bad at one thing, it does not guarantee that the person will also be good or bad at other things. An ability (IQ) test does not measure the same thing as an achievement test. Common statistical methodology, utilizing concurrent validity coefficients, demonstrates that a child obtaining an extreme IQ score (e.g., 130) would be expected to obtain a mean achievement level of only about 120. Thus the expected achievement of the child with an IQ of 130 is not 130 at all, but more likely to be about 120. For child with a low IQ score, the reverse is true: the expected achievement of a child with an IQ of 70 would be approximately an achievement score of 80. If practitioners do not adjust (regress) the scores they obtain in determining the "expected" achievement score, they will produce over-identification of high IQ children and under-identification of low IQ children. Without accounting for the correlation between two tests, one would never be able to know what the "expected" score should be.

The best way to determine the correlation between two measures is to look the correlation up in the manual of the test used. Unfortunately, not all manuals offer that information; the information is often based on absurdly small samples; and not all tests have been compared to each other. One way around this shortcoming is to estimate the correlation between the two tests. If we know the reliability of the ability test and the reliability of the achievement test, the correlation between the two tests can be estimated.

The equation for estimating the correlation between two tests is:

\[ r_{xy} = \frac{r_{xx} \cdot r_{yy} \cdot r_{xx} \cdot r_{yy}}{\sqrt{r_{xx} \cdot r_{yy}}} \]

**The values rxx and ryy are the internal consistency reliability coefficients for the aptitude and achievement tests used.

**Using the Tables to Determine the Expected Achievement**

The determination of expected achievement involves 2 basic steps:
When determining predicted achievement

1. Determine the estimated correlation between tests

2. Determine the expected achievement score

Go to Tables for Expected Achievement

Go to Determination of Severe Discrepancy

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Severe Discrepancy Determination by Formula

Kate obtains an IQ score of 90 and an achievement score of 74. Is this 16-point difference large enough to be considered a 'significant difference' between ability and achievement? Below is a table showing a statistical manipulation of the scores.

<table>
<thead>
<tr>
<th>Ability Score</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability of Ability Score</td>
<td>.91</td>
</tr>
<tr>
<td>Achievement Score</td>
<td>74</td>
</tr>
<tr>
<td>Achievement Reliability</td>
<td>.91</td>
</tr>
<tr>
<td>Correlation Between Ability and Achievement Scores***</td>
<td>.47</td>
</tr>
<tr>
<td>Predicted Achievement Score</td>
<td>95</td>
</tr>
<tr>
<td>Difference between Predicted and Actual Achievement</td>
<td>21</td>
</tr>
<tr>
<td>Magnitude of Difference required at .05 level</td>
<td>17.75</td>
</tr>
</tbody>
</table>

The 21.3 point difference between ability and achievement was found to be significant using the Predicted-Achievement method.

Using this approach to assessment, Kate should be considered to be functioning significantly below what was expected of her.

Above is an example of the use of a regression equation to determine a severe discrepancy. Regression is necessary because of the imperfect correlations between the ability and achievement measures. Just because someone has a low IQ score, we should not assume that they would also have a correspondingly low score on other measures. No one, I hope, would assume that just because a person had a very high IQ that the person would obviously be able to excel at gymnastics, or be able to hold their breath longer than 'normal.' Just because someone is good or bad at one thing does not guarantee that the person will also be good or bad at other things. The IQ test does not measure the same thing as an achievement test. Regressing the scores on the tests used also allows one to compare individuals to others with the same IQ. For example, if one child has an IQ of 120, while another an IQ of 80, the expectation of how each might be performing on an achievement test would be different. If the tests correlated at the .60 level, the first child would be expected to obtain an achievement score of 112 while the second child would be expected to achieve a score of 88.

Because most tests do not report scores in terms of Z, we can use a formula to calculated Z:

\[
Z = \frac{\text{Score} - \text{Mean}}{\text{Standard Deviation}}
\]
The simplest way to regress a particular score (either IQ or achievement) is to multiply the correlation between the two measures by the Z score you wish to regress. The formula would be:

\[
\text{Regressed score} = (\text{Correlation} \times Z)
\]

The best way to determine the correlation between two measures is to look the correlation up in the manual of the test used. Unfortunately, not all manuals offer that information; the information is often based on absurdly small samples; and not all tests have been compared to each other. One way around this shortcoming is to estimate the correlation between the two tests. If we know the reliability of the ability test and the reliability of the achievement test, the correlation between the two tests can be estimated.

The equation for estimating the correlation between two tests is

\[
0.7071 \times \sqrt{r_{xx} \times r_{yy}}
\]

**The values rxx and ryy are the internal consistency reliability coefficients for the aptitude and achievement tests used.

*** note that in the example at the beginning of this page, the correlation is listed as .47.

If the user knows from some source what the correlation really is, that should override the estimated correlation obtained from a formula. The formula provides the upper end of what is possible given the two reliabilities but those are estimates. In the example given, since the correlation between the IQ and the achievement (.47) was known, it was used. The example is something like using the PIQ versus a reading test (correlation .47) versus using the VIQ versus a reading test (correlation .65).

Use this link for Tables for the Reliability

Finally, the score must be turned back into a Standard Score with a mean of 100 and a standard deviation of 15. This is done by multiplying the "regressed Z" by the standard deviation and adding 100 to the result.

\[
\text{Predicted achievement score} = 15 \times (\text{Correlation} \times \left(\frac{\text{Score} - \text{Mean}}{\text{Standard Deviation}}\right)) + 100
\]

This regressed score provides the examiner with a different point for comparison. Comparing the expected score with the actual score gives a better idea of the magnitude of the difference between the scores obtained.
point difference was considered non-significant. However, she was expected (using the regression formula) to obtain an achievement score of 95, not 90. This increases the simple difference from 16 points to 21 points. We still don't know if this new 21-point difference is large enough to be considered a 'severe discrepancy.' For that determination, a different formula is needed.

Magnitude of Difference required at .05 level =

\[ 1.96 \times 15 \times (\sqrt{1-(\text{Correlation}^2)}) - 1.65 \times \left( 15 \sqrt{1-(\text{Correlation}^2)} \times \sqrt{1-(\text{RAbility} + (\text{RAchievement} \times \text{Correlation}^2) - (2 \times \text{Correlation}^2))/(1-(\text{Correlation}^2))-1.65 \right) \]

Thank heavens for computers!!

The first part of this monstrous equation:

\[ 1.96 \times 15 \times (\sqrt{1-(\text{Correlation}^2)}) \]

sets the level of our decision making to the 95% confidence (z=1.96) and then determines the Standard Error of Estimate for the obtained score. The second part of the equation:

\[ -1.65 \times \left( 15 \sqrt{1-(\text{Correlation}^2)} \times \sqrt{1-(\text{RAbility} + (\text{RAchievement} \times \text{Correlation}^2) - (2 \times \text{Correlation}^2))/(1-(\text{Correlation}^2))-1.65 \right) \]

reduces the final cut-off score by subtracting the standard error of the relevant difference score. The end result is a statistically justifiable 'severe discrepancy' upon which to make clinical decisions.

When a psychologist doesn't bother to struggle with the true identification problems, the result can be an over-identification of children as disabled when they are not. This increases the caseloads of the school and the individual staff, as well as making the whole label of "Learning Disability" into a meaningless catch-all. It must also be noted that even when a child does have a severe discrepancy between ability and achievement, this by itself does not constitute a diagnosis. It is a necessary, but insufficient factor, to be used in the determination.

For a wonderful explanation of this and other issues see:


A copy of the template for determining severe discrepancy using this method is available.

To download template press here.

Please send me an email if you download the template. I will be happy to answer any questions you may have. Please do not distribute the template to anyone else.
Subsequent to a discussion among several members of this list (Matthew Warren, John Willis, Ron Dumont, etc.), I decided there may be some confusion about regression methods used to identify severe discrepancies. I decided to write a more or less complete statement of those methods. I do hope it will be of value to someone. At the outset, I would like to thank John Willis, Ron Dumont, and Matthew Warren for their input. These are classy folks, as helpful as they are knowledgeable.

I will here try to describe the rational and method of using regression to determine severe discrepancies in diagnosing learning disabilities. A severe discrepancy in achievement occurs when a child's achievement deviates severely from what one would expect. It is essential, therefore, to establish expectation for a particular child. Few test scores ever coincide exactly with what is expected. In making a decision to label one discrepancy as severe and one as normal, some criterion must be established to which to compare an actual discrepancy. This application of decision theory will also be discussed here. I will then compare the method presented here with a method described by Cecil Reynolds in Chapter 24 of Handbook of Psychological and Educational Assessment of Children: Personality, Behavior, and Context by Cecil R. Reynolds (Editor), Randy W. Kamphaus (Editor). Hardcover (1990).

Terminology:

Y = Achievement score,
X = IQ score,
Y' = Predicted achievement score,
My = Mean achievement score,
Mx = Mean IQ score,
SDY = Standard Deviation for achievement scores,
SDx = Standard Deviation for IQ scores,
rYY = Reliability for achievement scores,
rXX = Reliability for achievement scores,
rXY = Correlation between achievement scores and IQ scores,
Ty = True score for achievement scores for a particular child,
Ey = Expected value of Y,
e = Y - Ey,
SE = Standard error of estimate when Y' is determined using X, and
zpn = normal deviate for probability = p and the number of type of test = n, one tailed or two.

There is no particular need to translate X and Y to the same metric. However, if this is done, it should be accomplished before calculations begin.

Assumptions:

1. Y is normally distributed,
2. The regression of Y on X is best described by a straight line,
3. Variance of Y on X is independent of X, and
4. The best method of determining EY is the method that minimizes SUM(e^2).

Because of assumption 2 above, the formula for predicting achievement given IQ is a special case of the general linear formula and is given by

Formula 1

\[ Y' = SD_Y(r_{XY}((X - M_X)/SD_X)) + M_Y \]

It can be shown that, given assumption 2, using Y' as EY will minimize SUM(e^2). Therefore, using Y' as expectation for Y will satisfy assumption 4 above, whereas using M_Y or X as the expected value of Y will not satisfy this basic assumption. One object in measurement is to minimize error. Since e is error, we would like to minimize it. However, since e is an unknown for a particular person on a particular administration of a test, we can only hope to minimize it within a group. Summing e across a group is fruitless. The mean is zero, and, therefore, so is the sum. If we square e before summing, then the result must be a nonnegative number contingent upon e. That is why we stipulate assumption 4 above. There are those who would like to use M_Y as an estimate of EY. Others would use X. Neither of these will minimize error. The attractiveness of either is based mostly on concern for a child not learning as well as his age mates and on the convenience of calculation.

While using Y' as an estimate of EY minimizes SUM(e^2) for a group, it may not minimize SUM(e^2) for a particular person. The task in the next section is to determine whether it is reasonable to believe that Y’ minimizes SUM(e^2) for a particular person. This is tantamount to asking whether Y’ = Ty.

To establish a criterion against which to compare actual performance, it is necessary to select a unit of measurement for deviations from expected. Given assumption 1 above, the natural unit of measurement is some type of standard deviation. In this case, SE is the appropriate unit. Given assumptions 1 and 3 above, SE is given by the following formula:

Formula 2

\[ SE = SD_Y(Sqrt(1 - r_{XY}^2)) \]
We next pose a question. The exact nature of the question reflects our philosophy of severe discrepancies. The two most common methods of stating this question are:

1. Is it reasonable to believe that, for child C, \( Ty = Y' \).
2. Is it reasonable to believe that, for child C, \( Ty > Y' \).

As Matthew pointed out, if we think we are concerned with question 1, then we would select a probability and normal deviate such that a deviation from expectation in either direction must be explained. Most school psychologists have tested children whose achievement scores significantly exceed expectation. This is sometimes more difficult to explain than the child who underachieves.

If we take the approach that severe discrepancies only fall below expectation, then question two is the appropriate question. Deciding which question is appropriate in a particular situation is of major importance. It is one of the two chief concerns in selecting a normal deviate for use in the next step.

Those trained in research will immediately recognize that the above questions correspond to the null hypotheses used in research. Question one evokes the use of a two-tailed test, while question two, a one-tailed test. In a very real sense, determining whether a particular child has a severe discrepancy is testing an hypothesis about that child. The logic is the same as in hypothesis testing. The null hypothesis is assumed to be true. This gives us a way of determining the probability of various events. We can tell which events are common and which are rare. When we test the hypothesis, we allow an event to occur and observe whether it is a rare event. The presence of rare events creates doubt about the truth value of the null hypothesis.

For example, suppose we are playing the old game, Twenty Questions, and are trying to identify an object. We have developed the null hypothesis that the object is a dog. Before venturing a "guess" as to what the object is, we propose to test the hypothesis with a question, the possible answers to which have known, relatively speaking, probabilities. We ask the question, "How many legs does this object have?" In my experience, the most likely answer, given that the hypothesis is true, is 4. However, in my life I have seen several three-legged dogs, one two-legged dog, and one five-legged dog (As an aside, I must admit that I paid 50¢ to see the five-legged dog). I have seen pictures of a six-legged dog and an eight-legged dog. Suppose we get this answer to our question, "It has six legs." This is not an impossible answer, but it is rare. It is so rare that most of us would decide to reject the hypothesis as untenable.

The question is, "How rare must an event be before we decide to reject the hypothesis as untenable in the face of the data?" As Reynolds (1990) argues, the traditional values are a likelihood of less than, or equal to, five in a hundred (.05 level), or less than, or equal to, 1 in a hundred (.01 level). Ultimately, the probability level must be set by the person making the decision.

Suppose we decide that any deviation from expectation, above or below, are of interest and that rare events have probabilities less than, or equal to, 0.05. On the normal curve, the normal deviate that corresponds to this decision is \( z = 1.96 \).
We would, therefore, calculate two critical values, one 1.96 SE above $Y'$ and one 1.96 SE below $Y'$. Actual values between these two critical values would be common events. Actual values outside these two critical values would be rare events. The formula for the critical values would be:

*Formula 3*

Critical values = $Y' \pm 1.96SE$.

If, on the other hand, we decide that only values below expectation are of interest, then on the normal curve, the normal deviate that corresponds to this decision is $z = 1.65$. We would, therefore, calculate only one critical values, 1.65 SE below $Y'$. Actual values equal to, or below, this critical value would be rare events. The formula for the critical value would be:

*Formula 4*

Critical value = $Y' - 1.65SE$

The two formulae may be generalized as follows:

*Formula 5*

Critical values = $Y' \pm z_{p2}SE$, and

*Formula 6*

Critical value = $Y' - z_{p1}SE$.

Matthew Warren posted some data to the list for which he had accomplished the calculations necessary to decide whether a particular child has a severe discrepancy. I have appropriated his data and will use it to illustrate the above formulae. The data are given below:

Matthew Warren wrote:

Scores:
FSIQ(wisc3) = 80
WJ(Writing Fluency) = 62
DATA:
Correlation (FSIQ, Writing Fluency) = .60
Reliability(FSIQ) = .95
Reliability(Writing Fluency) = .95

Calculated values:
Predicted WJ (Writing Fluency) = 88
Standard Error of Estimate = 12

Critical values (95% confidence), given that any deviation from expectation must be explained, would be an achievement score less than or equal to 64 or a score greater or equal to 112.

Critical value (95% confidence), given that only negative deviations from expectation are of interest, would be an achievement score 68 and below

Formula 1

\[ Y' = \text{SD}_y (r_{XY} ((X - M_X)/\text{SD}_X)) + M_Y \]

\[ = 15(.60((80 - 100)/15)) + 100 = 88. \]

Formula 2

\[ \text{SE} = \text{SD}_y (\sqrt{1 - r_{XY}^2}) \]

\[ = 15(\sqrt{1 - .60^2}) = 12. \]

Formula 5

Critical values = Y' +/- z(.05)2SE

\[ = 88 + 1.96(12) = 111.52, \text{ and} \]

\[ = 88 - 1.96(12) = 64.45. \]

Note that in the application of Formula 5, the first value, if not an integer, always rounds up to the next possible score, and the second value, down. In this case, rounding goes to the nearest integer, but that is not always the case. Therefore, as Matt said, test scores of 112 and above and 64 and below indicate a severe discrepancy at the .05 level of significance. If you have a machine that yields anything else, the machine is wrong.
Formula 6

Critical value = Y' - z(.05)1SE

= 88 - 1.65(12) = 68.2

Note that in the application of Formula 6, the value, if not an integer, always rounds down to the next possible score. In this case, rounding goes to the nearest integer, but that is not always the case. Therefore, as Matt said, test scores of 68 and below indicate a severe discrepancy at the .05 level of significance. If you have a program that yields anything else, it is wrong.

Up to this point, the formulae developed by Cecil Reynolds parallel the ones presented here. Cecil, however, at this point shifts his focus. When we test to see if a particular child has a severe discrepancy, there are four possible outcomes: 1) We correctly identify a child who really has a severe discrepancy [True positive]; 2) we correctly identify a child as not having a severe discrepancy [True negative]; 3) we erroneously identify a child as having a severe discrepancy when in fact he does not [False positive]; and 4) we erroneously fail to identify a child as having a severe discrepancy, when in fact he does [False negative]. The significance level, often called alpha, that we use, .05 above, is the probability of a false positive. The probability of a false negative is often called beta. The relationship between alpha and beta is inverse and nonlinear. If we decrease the likelihood of one type of error, then we increase the likelihood of the other. After developing all the formulas given above, Cecil decided that he should add something to reduce the likelihood of false negatives, the value of beta. He decided to do this without any notion what the value of beta was. He reduced the difference between Y' and the critical value of a two-tailed test by 1.65SEresid, where SEresid was defined in Critical Measurement Issues in Learning Disabilities in the Journal of Special Education, 18, 451-467, 1984.

For the current example, the critical value becomes 70.90. Clearly this does reduce the probability of a false negative, but it also increases the probability of a false positive. We are no longer working at the .05 level, but at the .1556 level. Cecil (1990) cites an example where the relevant z-score moved from 2.00 to 1.393. This changed the probability from about .05 to .1646. He seemed to have been somewhat confused about the question he was asking at the time. He changed the probability to .082, as it would have been in a one-tailed test. He clearly started his discussion using a two-tailed test, then stated that severe discrepancies went in only one direction. When he reduced the distance to the critical value, he subtracted a value based on a one-tailed test. The situation then is this: he selects only one of the two critical values from a two-tailed test and used a value from a one-tailed test to move that toward the mean. Confused? Worry not; it gets worse. Cecil then drew a picture (his Figure 24.2) to clarify matters. In this figure, he shows only one of the critical values of the original one-tailed test (happens not to be the one discussed in the text). Then he both subtracts 1.65 and adds 1.65 to this critical value to get two more critical values. Thus, he takes what should be a one-tailed test at the .05 level, stacks it on top of a two-tailed test at the .05 level, but runs it in both directions so that the probability would be .10. At this point, I think I will give up trying to explain what he proposed. The logic is clearly muddled and the issue of significance level becomes totally garbled. In the two examples that I ran, Cecil's and Matthew Warren's, the significance level multiplied by about 3. I think there is no way to predetermine what will happen to the significance level, but clearly it alters drastically with the addition of Cecil's invention.

Interestingly, Cecil added the "correction" to control beta. He describes no way of determining what beta is, either before or after his fix. There are methods of controlling beta, but not with Cecil's formula.
Here's the damage done by this method. Cecil gives a lengthy discussion of why the .05 level of significance is appropriate. However, when he acknowledged that the significance level changed, he changes terminology. Instead of significance, it becomes the "percent of the total population." Would an astute reader miss this shift? Ron Dumont and John Willis, the best in the business, missed it. On their template for calculating severe discrepancies using Cecil's method, they specified that the results are at the .05 level. Has anyone else missed it? In the WIAT manual, page 188, the significance level is clearly identified as either .05 or .01, when it clearly is not. Big Cecil himself acknowledged that the chances changed (1990, p. 552).

Developing procedures to assume control of beta, correctly is beyond the scope of this post. I may do that another time.

Also, the WIAT manual (p. 189) implies that Cecil's procedures were used to establish the significance bands in its tables. I have neither the time nor inclination to check that assertion. Certainly I would view those tables with suspicion.

Hubert

Irish Blessing
May the road rise to meet you.
May the wind be always at your back.
May the sun shine warm upon your face.
And rains fall soft upon your fields.
And until we meet again,
May God hold you in the hollow of His hand.

For a download of a severe discrepancy analyzer that follow Hubert's description press here.
Does Subtest Order Make a Difference?

The following are exchanges from the Institute for Applied Psychometrics website and discussion group focusing on the Cattell-Horn-Carroll Theory of Cognitive Abilities located at http://home.att.net/~gfgc/index.htm

Participants in this particular discussion include: Ron Dumont; Laurie Ford; Richard Woodcock; Mark Daniel; Samuel O. Ortiz; Kevin McGrew; John Willis, Catherine A. Fiorello

Kevin and Dick

I am hoping you can respond to a question that comes up every time we do a workshop on the Cross battery approach. Here it is paraphrased:

"Given that IDEA requires that any standardized tests used "are administered in accordance with any instructions provided by the producer of such tests;" can the WJ-R subtests be administered out of sequence?"

Have I missed anything in the WJ-R's manual that tells the examiner that it is okay to pick and choose subtests as they are needed? If the WJ-R is used to supplement other cognitive measures (a la Cross battery) does that go against the IDEA conception of a safeguard.

Are there any studies available to tell us that giving subtests out of order, or out of the framework of the standardization procedure will, or in this case, will not have an effect on the results? We have always heard about the need to follow standardized procedures so as not to artificially introduce confounding variables to the results. I think in particular about some of the memory tasks. Visual Auditory Learning is subtest #8 in 'regular sequence'. Does administering it as subtest #2 have any potential impact on the results?

I hope this question makes sense. If not, blame it on John!

Thanks

Ron

Ok, I am not Kevin or Dick but having worked with the two of them for years and having done more than my share of WJ trainings over the years I think I can help you on this one....
The key here lies in part in the issue of calling them "Tests vs subtests".....the WJ system of "Tests" (not subtests) is developed so that each individual test (e.g. Memory for Names, Sound Blending, Calculation) is a stand alone test by itself.....Dick often uses the analogy of a tool box which I like. You reach into the box and pull out the tool you need to help with your referral question....with this in mind you really would not necessarily have to use all the tests in the battery so they have to stand by themself (different from the subtest format of the Wechsler scales). This is where the selective testing table in the manual is helpful. With that in mind (you could use many different combinations of tests) ordering of the tests from a psychometric point of view is really not an issue, from a convenience or preference of the child, age of the child, etc....that may be where you address ordering......the test was developed so each test stands alone as an individual" test " in a battery of tests"....some of this is discussed in the first few pages of the technical manual (pp 4-7). Of course ya gotta give Memory for Names and VAL before MforN and VAL Delayed Recall (ha!....WJ humor)....

There is some logic to the ordering of the tests in the easel booklets but as you can see they are in a binder so you can reorder them and repackage them to meet your needs. In my trainings I have found people who work with younger children find Memory for Names a great place to start while those working with secondary age students and adults do not like to start with Memory for Names....so you can change it up.....this is an especially nice feature when you are using a cross-battery approach like Kevin and Dawn have proposed......

Long answer, good question, common question.....hope this information helps with your work,

trainings, etc....

Laurie

---

Ron, you will find an example of such a statement in the middle of p. 40 of the COG Examiner's Manual. In some places where we make this statement we also point out that the tests were not presented in any particular order during standardization.

I am not aware of any studies that say there is, or isn't, such an effect. I doubt that there could be anything more than a trivial effect. Test developers expect (or, at least, hope) that tests are independent and robust enough that the order of administration is not important in a situation where the rapport is good and the subject is not tired. Of course, if there is something learned in one test that will help the subject in another test, then order would become quite important.

Richard Woodcock

---

I don't know of published research on this question, but have some anecdotal information. I once worked for an organization that provides batteries of individually-administered tests to teenagers and adults for vocational and educational counseling. The battery sequence was changed from time to time. There was a very noticeable depression of performance on the first one or two subtests in the battery. As I recall, the magnitude of the effect was about a 10 percentile point shift in the median (i.e., when the subtest was given first the average performance was at about the 40th percentile on the norms based on later administrations).

It's good to see that the WJ-R standardization utilized varied administration sequences, so that this effect would have minimal influence on the norms. A similar approach was taken with the diagnostic subtests on the DAS.

Mark Daniel
Although others have already offered their perspective on the issue raised by Ron with respect to the WJR in particular, I would like to add some comments regarding the legalities Ron raises as they pertain to any cognitive test used under the auspices of IDEA.

The concern with the legal issues raised by the questions posed to Ron is in reference to Sec. 300.532 Evaluation procedures of IDEA, which states in part:

(c)(1) Any standardized tests that are given to a child—
   (i) Have been validated for the specific purpose for which they are used; and
   (ii) Are administered by trained and knowledgeable personnel in accordance with any instructions provided by the producer of the tests.

First, I think it is important to note that the law will always be wholly unable to keep up with many of the advancements being made in the science of assessment; cross-battery assessment being only one such advancement. The notion of ability-achievement discrepancy, no matter how discredited it continues to be in the literature, nevertheless remains a central criterion for determining the presence of a specific learning disability and represents another problem inherent in the "practice-legal gap." Apologies to Kevin for jumping on his "theory-practice gap" idea, but the notion is the same. As such, we must often decide whether to operate strictly within the legal mandates as written or apply superior methodology when and where available. For my money, in a court of law, I would rather sit on strong science and reason in keeping with the spirit of the law than attempt to justify my actions under the letter of the law. Legal mandates are not, in fact, straightforward, rigid, or inflexible. The actual meaning of any statute tends to be fluid and susceptible to myriad interpretations as evidenced by the many, often conflicting, decisions rendered in courtrooms across this nation.

Now if we extend the question posed by Ron into actual practice and we adhere to it as written, are we not then bound to administer every single subtest (in the correct order, of course) from a particular battery? In other words, in the absence of explicit instructions from the publisher that allow for the administration of only specific subtests, are we not legally required to administer the entire test, even if we may be only interested in a particular set of tests or composite score? It seems to me that in a wide variety of cases, individuals have been given only a portion of certain batteries because other portions were deemed invalid or inappropriate for that individual. In the case of children who are culturally or linguistically diverse, giving only the Performance subtests of the Wechslers has been a common practice for decades and remains so even today. Similar selective administrations are often done with children who are blind, deaf, motor impaired, etc. Is it illegal to do this then, because the verbal tests were not intermixed with the nonverbal tests and vice versa? Similarly, what if I only want the non-verbal composite from the K-ABC because I deem that composite to be the only appropriate measure of ability for the individual in question? Must I give all of the K-ABC subtests and then only score the ones loading on the non-verbal scale? The correct "legal" answer to all of these questions might well be "yes," but how many psychologists adhere that strictly to the letter of the law? Likewise, I have yet to read in any publishers test manual that the only valid use of the test occurs only when every single subtest is given. Rather, test publishers are savvy enough to recognize the potentially lucrative alternative uses of the test and rarely do anything to discourage exploration or adoption of such use. In fact, when it is backed by empirical research, it appears that test developers actually encourage such use, albeit it is not always specifically endorsed in the test manual. Thus, test developers seem to prescribe mainly the manner in which the entire test is to be given in order to derive certain composite scores. They do not often, however, offer prescriptions prohibiting the use of the test in alternative manners. Doing so would likely amount to economic suicide.

The discussion above addresses the issue of giving subtests "out of the framework of the standardization procedure." I'm sure I've made it clear that not only do test publishers routinely omit prescriptions against using selected subtests or portions of the test, they also do not state that the use of the battery and interpretation of resulting scores is valid only if all subtests are administered. It seems to me that such alternative uses (whether with particular populations or through different theoretical foundations) is left up to the professional judgment of the examiner and that the examiner, in such cases, would be responsible and liable for providing a suitable rationale for whatever decisions were made and action taken.

As for giving subtests "out of order," I am certain that it does indeed violate standardization for some tests (but not the WJR or DAS according to Dick and Mark). I am not certain, however, that violation of a particular administration sequence produces noticeable effects one way or the other (Mark seems to say yes, Dick seems to say no) so I'll leave that up to them to debate. I do know that subtests are generally given in a particular order to attend primarily to issues of examinee interest, rapport building, minimization of fatigue, counterbalancing, etc., as well as maintaining consistency in the experimental (test) setting. If sequence is deemed important, for whatever reason and for whatever test, use of cross-battery assessment does not require that tests be given "out of order" or out of sequence per se.
Based on the preceding discussion the order of subtests seems to matter most when the examiner is interested in using the full battery and deriving a full scale score. If the examiner is not interested in a full scale score, then I see no reason why the examiner can't give only the selected subtests of interest, and no reason why the examiner can't give those subtests in exactly the order they occur within the battery as a whole and exactly in accordance with the instructions provided by the publisher for administration. The WJR is a good example that allows such an approach, albeit many batteries also offer a variety of composite scores that can be calculated on the administration of certain specific subtests. For example, if I am interested in assessing an individual's Reading Aptitude or Oral Language using the WJR, the manual specifies exactly which subtests must be administered. Although presentation order is not specified, there is no reason to give subtests 2, 3, 11, and 13 (for Reading Aptitude) or subtests 2, 6, 13, 20, and 21 (for Oral Language), "out of order." Subtest 2 can always be given before subtest 3 and, subtest 3 can always be given before subtest 11, etc. I see no reason why derivation of Gf, Gc, Gs, etc., clusters is any different (especially on the WJR which is constructed specifically on these concepts) and thus, interest in a particular cluster or composite does not have to violate sequence. Imagine having to give ALL 21 subtests of the WJR in order to derive a valid score for the Oral Language cluster! Surely, no one would bother if that were the case and thankfully, Richard was careful to make it clear that this was not necessary. Thus, cross-battery assessment procedures do not necessarily have to violate order or sequence of subtests and neither do they violate administration procedures.

Although cross-battery techniques are based on an alternative theoretical framework, administration and scoring issues remain well in line with the "intended" use of any battery which is to measure or sample a particular aspect of ability or behavior. Because the administration and scoring procedures specified in cross-battery assessment are entirely consistent with the procedures specified in the test manuals, cross-battery assessment does not explicitly violate any instructions set forth by the publisher. In cross-battery assessment, it is only the combinations of subtests that are used which may not be described by the tests' manuals, but which are supported by the empirical research undergirding the entire cross-battery approach. In effect, as Kevin is fond of saying, taken together the research behind the cross-battery approach and the methods specified by it, comprise a full fledged technical manual which guides how tests are to be combined and interpreted as much as is found in any other test manual. Therefore, as long as cross-battery assessment is carried out as specified, an examiner could further argue that he or she is indeed following the instructions set forth by the developers of the "test." And finally, because cross-battery assessment is much more than simply "picking and choosing" subtests (precisely because it is based on empirical research and reasoning, not haphazard guessing), I see nothing that constitutes any inherent violation of the IDEA regulations cited at the outset. In fact, given that cross-battery assessment can be used to generate a test battery that is unique to the individual's needs and specific to the referral concerns, I see a much more solid foundation for assessment and an approach that is highly defensible in a court of law.

Sorry once again for the mini "dissertation" but it seemed relevant. And I don't want to hear anything about my long winded tendencies from either of you, Kevin and Dawn :)

Sam

Sam the "orator" Ortiz has struck again.

I'll be brief, non-technical and non-legal.

When I take my seriously ill child to the doctor to help with a tricky diagnosis I'm more concerned about the doctor blending his/her clinical skills with his/her instrumentation, even if it means that he/she may modify a procedure to meet my child's unique needs. There is both an "art" and a "science" in psychological assessment and interpretation. This is one of the reasons I left the school psych. trenches after 10 years. I felt I was no longer a school "psychologist" but had become a school "proceduralist" because rules and regulations where suppressing my art variance. I concur with Sam that it is better to do what is in the professional best interests of a child during an assessment and wage war with administrators, regulations, etc. instead of letting these forces wage war against good practice.

There --- I've been waiting 9 years to get that off my chest.

Kevin McGrew
Thanks to everyone who has responded. Let me add another kink. I do believe that administration order can effect the end results, similar to what Mark talked about. Let me explain:

I was asked to help norm a task (the Mesulam) of visual scanning being used by some evaluators. The Mesulam Continuous Performance Test consists of two pages with the letters of the alphabet printed in uppercase. On one page (Ordered), the letters are placed in neat, orderly rows and columns, while on the second page (Random), the letters are placed in a haphazard fashion, with no order imposed. On both pages, 60 A’s are placed among the other letters. Regardless of the page, the A’s are in the same location, with approximately 15 in each of the 4 quadrants. Because there were no administration directions and because some psychologists reported giving it differently (one used the ordered page first, the second gave the random page first) we decided to give the task both ways (approximately half got the ordered page first, half random page first). Because the total score (score on Ordered plus Random) was the score of importance, few thought that the administration order mattered. The Mesulam was administered in a balanced manner, with 849 children being administered the Ordered page first (Ordered Administration) while 691 were administered the Random page first (Random Administration).

Here is a snip of our results:

Means were computed for each of the nine age groups (6 to 14), and for both administration procedures (Ordered and Random). Analysis was done to determine the effects of sex, age, and order of administration. No significant differences were found between the sex of the subjects. In contrast, both order of administration and age were highly significant. Effect of administration order and age was assessed by a two factor analysis of variance. OE was found to be significantly affected by both the administration order (F=6.709, P=.0097) and the age level (F=14.654, P=.0001). RE and TE were found to be non-significantly affected by the administration order (RE: F=.098, P=.7652, TE: F=2.934, P=.0869), but significant for age level (RE: F=10.379, P=.0001, TE: F=18.045, P=.0001).

For ages 6, 7, and 8 there was a pronounced difference in the total number of errors (TE) made depending on the order in which the pages were administered. For example, six-year-olds given Ordered Administration had 5.4 total errors while those six-year-olds given Random Administration had an average of only 2.7 total errors. Interestingly, most errors occur on the first page presented (Ordered or Random). Whichever page was administered to the child first yielded comparatively more errors than the succeeding page. Six-year-olds given Ordered Administration averaged 3.6 errors on the Ordered page, compared to the 1.8 errors on the Random page. Similarly, six-year-olds given Random Administration averaged 1.9 errors on the Random page, while averaging .74 errors on the Ordered page.

Maybe it had something to do with this specific task, but it sure raises the question of unanticipated administration influence.

Ron

Just to add to the confusion (it's what I do; it's a gift), I had a colleague who became tired of low scores for Coding on the WISC-R, so -- anticipating the WISC-III -- my colleague began giving Coding early, rather than late in the sequence of subtests. Lo and behold, the Coding scores of the post-change evaluations were generally higher than the Coding scores on the pre-change evaluations.

As Sam and Kevin recommend, I do use subtests out of context. In addition to the non-random McGrew, Flanagan, and Ortiz Integrated Carroll/Horn-Cattell Gf-Gc Cross-Battery Approach, I sometimes have a specific skill I want to study, for example using the KAIT listening comprehension subtest for a student who is having difficulty with high school lectures. I will even intersperse other tests of approximately the right length and then administer the listening comprehension delayed
Does Subtest Order Make a Difference

recall! However, I worry -- not about the numerical sequence with or without gaps -- but about the effect on test "n" not being preceded by tests "n-1" and "n-2," as with Ron's Mesulam or my colleague's WISC-R Coding. It is not preserving the order 4, 11 (vs. 11, 4) to assess Ga on the WJ-R that concerns me. It is the possible effect of 4 not following 3, or 11 not following 10, or 11 not coming fairly late in the assessment session. Obviously, Sam is correct that I would be insane to give the entire WISC-III and then give the entire WJ-R just to get at tests 7 and 14 for Gf. In practice I do what Sam and Kevin recommend, but I do worry.

Knowing that the WJ-R tests and DAS Diagnostic Subtests were actually normed in varied sequences assuages some of my guilt when using those instruments. Ron's Mesulam effect and Mark's experience with his test would, however, still be a concern. I don't know for sure, for instance, whether recognizing Jawf and Meegoy might be equally difficult at all times during a 90-minute test session but remembering that all kinds of life live in the sea might become much more difficult toward the end of a long test session than at the beginning. Or vice versa -- who knows? If the differences are small, who cares? But what if there is some test or subtest for which the difference is significant?

I seem to recall that when the WISC and WISC-R were being compared, it was found that, as expected, students had higher WISC than WISC-R scores. However, if the WISC were given first, the difference was much smaller than if the WISC-R were given first. Apparently the more liberal querying rules on the WISC-R gave students a set for providing more elaborate, higher-scoring answers on the WISC if they took the WISC second.

I absolutely agree that I must practice my art to the best of my knowledge and ability and not be a proceduralist [McGrew, K. S. (1994). School psychologists vs. school proceduralists: A reply to Willis and Dumont. Communique, 22 (8)] and I share Sam's belief that I can explain and defend my science on the stand. My concern is that my science might be flawed in the case of certain tests or subtests given out of their norming-sample context and I don't know which ones (if any) they might be.

John Willis

It's not so much that I or Kevin recommend that this be done, rather it is more that using subtests outside the context of the whole or entire battery does not appear to constitute any legal transgression as defined under IDEA. Moreover, cross-battery assessment is not the first or only method that takes subtests out of context of their corresponding battery. As I mentioned in my previous post, test context is violated in any number of ways when working with individuals from various populations (culturally/linguistically diverse, deaf, blind, motor impaired, etc.) or whenever the entire set of subtests are not given. Standard methodology in neuropsychological evaluation has long used subtests out of context without any seeming legal or scientific ramifications. If such methods have passed both legal and scientific scrutiny currently and in the past, why would cross-battery methods be judged any differently? Now whether there truly is or isn't any scientific reason against the use of subtests out of context is something that appears to remain a point of debate. In my opinion, and I'm sure Kevin and Dawn as well as others will share this view, the context of a test battery that is either unfounded or poorly founded upon theory is less defensible from an empirical standpoint next to a context that is steeped in theory and backed by research. Thus, one of the core arguments underlying cross-battery assessment is that it is deliberately designed to provide a better, more defensible context than that provided by the subtests' battery itself.

I believe that observations about performance on any given test, at any point in the assessment process, is a major responsibility of the assessor whether they are using cross-battery techniques or not. If it is noted that an individual is tiring heavily under the burden of testing, the examiner must be aware of it and attend to it. Is the validity of an individual's score on WJR Concept Formation any higher simply because it was given after the preceding 13 subtests even though the individual was noticeably fatigued vs. giving it along with subtest 7 (Analysis-Synthesis) without the preceding/intervening tests on a day when the individual was fully rested? Clearly, some issues that may be related to sequence or order are within the control of examiner and some are not. Given the lengthy history of psychometrics, it would seem to me that if sequence and order were of such critical importance, test publishers would long ago have constrained administration to very rigid standards and procedures. It would make sense that prohibitions against altering the sequence and context (administering every and all subtests each time a battery is used) would be explicitly stated and repeated in the manual of every major test battery by now. Again, from a practical point of view, forcing professionals to use batteries
Does Subtest Order Make a Difference

in such a manner would also surely doom a test's acceptance and usage. I would like to think, based on Dick's and Mark's comments, that developers have looked at the issue and that either it makes little difference or they've made it a non-issue by simply standardizing the test with various sequences.

I don't know if it will help your angst much, John, but I would venture to guess that the tradeoff may well be an acceptable one. That is, although we may not know which subtests might be significantly affected by being given out of the context of the test as a whole or even out of sequence, the advanced and defensible theoretical context created by cross-battery assessment methods balances that and any other concerns that accompany the application of the method and appears to be of equal if not greater value than the alternative. The growing recognition in psychometrics of the necessity for "theory first" will likely hold those in good stead whose practices are consonant with the credo for some time to come.

Sam

It appears that in some specific cases order of administration matters. For example, Ron showed that order of administration mattered on the Mesulam Continuous Performance Test. This is not surprising given that the demands of the task were essentially identical for both pages. It was also not surprising that fewer errors were found on the second page (whether is was the Order or Random page) because practice effects were likely operating. Although many practitioners would recognize that order of administration would likely have an effect (i.e., better performance on page 2) on tests that are similar to the Mesulam test, other instances in which practice effects or other variables may be operating with regard to administration order may not be obvious at all. Notwithstanding, it is important to keep in mind that the cross-battery approach espoused by Flanagan, McGrew, & Ortiz does not recommend the interpretation of individual subtests. Rather, their cross-battery approach recommends the interpretation of broad ability clusters (i.e., the aggregate of qualitatively different indicators of the broad ability). In instances in which significant differences among subtests are found within broad ability areas, the cross-battery approach recommends that additional measures of the narrow abilities underling these subtests be administered to determine if there are differences between narrow ability clusters within broad ability domains. Thus, interpretations within the context of the cross-battery approach are always made based on clusters consisting of individual subtest scores that vary consistently.

This type of approach to interpretation is sensitive to tests that deviate from tests with which they should be consistent theoretically (and empirically). Therefore, if order of administration had an impact on a particular subtest (either positively or negatively), it would most likely be inconsistent with other tests administered that measure aspects of the same broad ability. The cross-battery approach does not recommend interpreting this finding in isolation. Other measures would be administered to determine if this finding is spurious or if it represents, for example, a meaningful and substantiated strength or weakness for the individual. In short, the guiding principles and procedures of the cross-battery approach are designed to ensure that anomalous findings are identified as such and interpreted appropriately. Furthermore, all interpretations within the context of the cross-battery approach (or any other approach for that matter) should be supported by additional sources of data. In my opinion, within the context of the cross-battery approach, order of subtests is a trivial matter. If subtest order made a difference either positively or negatively for a given individual, then this would most likely be detected and interpreted as such in an appropriately designed cross-battery assessment.

Dawn Flanagan

My best guess as a practitioner coincides with your better-informed judgment (for which I am grateful). I began switching from teaching to testing in a rehabilitation center where I seldom encountered a student or adult client who could take a standard test battery in toto and before that had the privilege of assisting the late Martin Berko in his assessments of children with cerebral palsy. However, it is the angst that drives my assessments, waking me up in the middle of the night wondering, "Did I do that right?" "Would that have made a difference?"

My favorite book (aside from Marty and Frances Berko's) on the importance of observations, as you observe, is O'Neill, A. M. (1995). Clinical inference: How to draw meaningful conclusions from tests. New York: Wiley. Paying attention to the student does cause problems. I have had assessments that stretched into a fortnight
Does Subtest Order Make a Difference of quarter-hour sessions.

John Willis

I agree with John's point that if there are large and systematic effects of administration order, we at least ought to know about them. To take Kevin's medical analogy, we'd expect our physician to be aware of any tendencies for environmental factors to skew the results of a diagnostic test. So, for example, if it were true that examinees perform substantially worse on the first one or two tests they take, this would be important information and would affect interpretation.

Judging from the discussion thus far, it appears that administration sequence is an under-researched topic. Publishers have long presumed that sequence may have an effect, which is why the manuals recommend that subtests in a battery be administered in a standard sequence. What publishers don't know is the actual effect of departing from the standard sequence, or what kinds of departures are most significant.

John also distinguishes two types of possible effect: (1) general location (early, where there may be "start-up" effects, or late, where there may be fatigue effects), and (2) sequence (the effect on Test A of having given Test B immediately before it). Regarding #2, in my experience test designers have generally tried to avoid giving tests back-to-back that are highly similar either in the ability they measure or in their task features. Giving highly similar tests back-to-back would seem to invite practice and/or fatigue, and might affect motivation (i.e., when a child feels they have done poorly on one test, how do they approach the next test if it appears the same?) These questions are relevant to the practice of cross-battery assessment. Should practitioners be urged to avoid giving several tests of the same ability consecutively?

Mark Daniel

To continue the discussion, Dawn noted the following about the results we got on the Mesualm task - the fact that there were differences on the pages based on the administration order: "This is not surprising given that the demands of the task were essentially identical for both pages. It was also not surprising that fewer errors were found on the second page (whether is was the Order or Random page) because practice effects were likely operating."

This was my point, albeit poorly said. Mark's follow up is important as well.

If we choose to test Processing Speed as an example, by giving just the 2 WJR subtests (visual matching and cross out) might they be affected by both Dawn's point (practice effect for very similar tasks) and by Mark's point (sequence effect). It is interesting, if true, that there is little empirical evidence to support or deny the effects. Sounds like a dissertation topic.

Ron

Thanks for your comments Mark. In response to your question, "Should practitioners be urged to avoid giving several tests of the same ability consecutively?" I believe the answer is yes. Upon designing a cross-battery assessment, practitioners should be cognizant of the demands of the tasks and the nature of the stimuli with regard to administration order for the reasons that many have offered in this dialogue. However, the cross-battery approach suggests that practitioners select a minimum of two subtests that measure qualitatively different aspects of each of the broad abilities they wish to assess. This guiding principle of the cross-battery
Does Subtest Order Make a Difference

approach may serve to minimize the effects of practice.

It would be interesting to hear from those who use the cross-battery approach or to examine the tasks that someone has selected recently for cross-battery assessment. I am in the process of having a student review a random sample of cross-battery evaluations from a referred sample of 200.

Dawn Flanagan

I try to use clinical judgment when arranging the order of subtests in a cross-battery assessment. As Mark suggests, I try to vary the abilities measured and the modalities. I keep time spans in mind for memory and such. I do often use Naglieri’s Planning subtests, and make sure I do those first so that the Coding/or whatever response set of left to right doesn't contaminate the Planned Codes freedom. But I agree that this is an under-researched question....

Catherine A. Fiorello, Ph.D.
Obtained Score Differences on the Mesulam Continuous Performance Test: A Comparison between ADHD Subjects and Controls

Ron Dumont, Casey Stevens, Margaret M. Dawson, Richard Guare, and Michael Weiler

Normative data for the Mesulam Continuous Performance Test were gathered from a sample of approximately 1500 controls aged 6-14. These data were used as the basis for comparing the Mesulam performance of 170 children independently identified as having an attention-deficit disorder (ADHD). Due to noted methodological shortcomings, this article serves merely to highlight differences in task performance between groups. As these differences were quite drastic, and thus compelling, research within a more rigorous framework is warranted. Should the Mesulam withstand closer scrutiny, it appears to have advantages over other, more expensive CPTs. It may be found useful as a screener in schools, administered to large groups of children at once, and as a useful tool to be employed in comprehensive psychoeducational evaluations.

Attention-deficit/hyperactivity disorder (ADHD), a disorder originating in childhood and characterized by inattention, impulsivity, and hyperactivity, is estimated to occur in 3 to 9 percent of the school-aged population (Szatmari, 1992), although it has been argued that rates might be much higher than commonly estimated (Shaywitz & Shaywitz, 1988). In recent years, the widespread recognition of the disorder by the popular press has led to increasing numbers of children referred for diagnosis. Along with this has come concern, among both professionals and the lay public, that the disorder is being over-diagnosed. Thus, there is a pressing need for accurate diagnostic procedures.

Diagnosing ADHD typically requires a medical and developmental history of the child along with parent and teacher ratings of behavior. While such information is essential, there is a subjective aspect to this data that can restrict its utility. Evaluators have desired to supplement this information with methods which enable the attention problems to be observed and quantified on a more objective basis. In an effort to satisfy this need, laboratory tasks have been employed in the assessment of attention disorders. For example, beginning in the 1970's, Sykes (Sykes et al., 1973) engaged subjects in laboratory tasks to demonstrate deficits in sustained attention and vigilance. The early laboratory tasks required subjects to attend to visual or auditory stimuli and respond differentially to target and non-target stimuli. Thus, for example, letters would be flashed sequentially on a screen and the child would be asked to press a button each time the letter X appeared when it was preceded by the letter A.

More recently, common laboratory measures of sustained attention have included computerized continuous performance measures such as the Test of Variables of Attention (TOVA, Greenberg, 1991), the Conners’ Continuous Performance Test (CPT, Conners, 1994), and the Intermediate Visual and Auditory Continuous Performance Test (IVA, Sanford & Turner, 1994). Although presentation formats and task demands are similar to earlier attention tasks, the computerized nature of these measures moves them out of the realm of research tasks requiring sophisticated data analysis skills and specialized equipment and makes them accessible to evaluators for use as part of their diagnostic procedures. However, these tasks typically require, in addition to computers, an initial investment in software and in some cases, a per test administration fee.

Reviews of research on computerized continuous performance tasks have generally been favorable, and they are seen as playing a role, albeit limited, in the evaluation of attention disorders. Barkley and Grodzinski (1994), for instance, evaluated the utility of neuropsychological measures, including continuous performance tests (CPTs) for distinguishing children with ADHD from normal controls and children with learning disabilities. They found CPT measures among the
most useful of the assessment procedures investigated. Nonetheless, they noted that positive but not negative CPT findings can have diagnostic utility. Thus, while poor performance on a CPT measure was indicative of an attention disorder, good performance did not necessarily rule out attention disorders. These results have been replicated by Matier-Sharma and her colleagues (Matier-Sharma et al, 1995). With the proliferation of research on attention and other executive skill tasks, there remains hope that some combination of tests and behavioral observation procedures will further refine the diagnosis.

Paper and pencil tasks have also been used with varying results to examine aspects of attention. Cancellation tasks, which require the individual to cross out a figure or letter from a visual array have been among the more popular paper and pencil tasks. Aman and Turbott (1986) found that a cancellation task discriminated between hyperactive and control subjects. Voeller and Heilman (1988), using a letter cancellation task, found that boys with attention disorders made more errors of omission than a group of normal controls. Weyandt and Willis (1994), however, did not find a significant difference between ADHD children and normal controls on a visual search task.

Should pencil-and-paper cancellation tasks prove effective in discriminating between ADHD and non-ADHD populations, they have several advantages over computerized measures. They are convenient, economical, and portable. Furthermore, they require less time to administer and score than do computerized measures.

As Barkley and others (1994) have noted, in order for a test to be diagnostically useful, it must be able to not only identify the children with ADHD, but it must also accurately identify children without ADHD.

Ellwood (1993) discusses parameters that can be used to examine a test’s diagnostic usefulness. Test specific parameters include sensitivity, or the proportion of individuals with a disorder that exhibit the sign (i.e., the proportion of children with ADHD who receive scores within the abnormal range) and specificity, or the proportion of individuals without a disorder that do not exhibit the sign (i.e., the proportion of controls who receive scores within the normal range). These two parameters are calculated in the research setting by first knowing the diagnosis of the children (through test-independent criteria) and noting how they perform on the test of interest. However, as Ellwood (1993) points out, this is the opposite of the way an evaluator uses a test. The evaluator starts with the test score and attempts to determine the child’s diagnosis. In order to judge the usefulness of a test for this purpose, the evaluator will need to look at a test’s sensitivity and specificity in light of the disorder’s base rate in their referral population.

For example, if a test was used as a screening measure on a population of 1000 children in which 4% (40) of the children have ADHD, and that test gives an abnormal score for 90% of the children with ADHD (i.e., sensitivity) and gives a normal score for 90% of the children without ADHD (specificity), the following diagnostic properties result.

Table 1. Calculation of Sensitivity, Specificity, PPP, and NPP

<table>
<thead>
<tr>
<th></th>
<th>ADHD</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal Score</td>
<td>a 36</td>
<td>b 96 132</td>
</tr>
<tr>
<td>Normal Score</td>
<td>c 4</td>
<td>d 864 868</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>960 1000</td>
</tr>
</tbody>
</table>

Sensitivity = a/a+c = .90
Using this table, one can calculate Positive Predictive Power (PPP), or the chances that a child who receives an abnormal test score actually has ADHD. PPP = a/a+b = 36/132 = 0.27. A test with 90% sensitivity and specificity has restricted usefulness as a diagnostic tool if it is used on a population with a 4% base rate of the disorder because if the child receives an abnormal score, (s)he is still much more likely to be a control than a child with ADHD.

There are two ways of making such a test more useful diagnostically: 1) Use it on a population with a higher base rate of the condition (e.g., the base rate of ADHD at an ADHD clinic is likely much higher than the 3-9% identified in the general population), and/or 2) increase the cutoff point for abnormal scores so that specificity is as high as possible.

Greenberg and Crosby (manuscript in preparation) applied the first two measures to the performance of persons diagnosed as ADHD with those of normal children on the Tests of Variables of Attention (TOVA). Using a cutoff set for a 20% false positive rate, the TOVA yielded sensitivity and specificity rates of .73. If this test were used to screen a 4% base rate population for ADHD, only 10% of the test positives would be true positives (PPP), while 97% of the children who received a normal score would be true negatives (NPP).

In addition to distinguishing between children with and without an attention disorder, an instrument that purports to be a valid measure of attention should also be sensitive to the developmental changes associated with attention. As with other cognitive traits such as memory, research suggests that as children age, their attentional capacity increases, and a variety of measures have been used to demonstrate this phenomenon.

Routh, Schroeder, and O’Tuama (1974) found that activity levels in children declined with age when measured by open-field activity ratings and parent-completed behavior checklists. Levy (1980) found age-related changes in sustained attention using a continuous performance test, in motor inhibition using a line drawing task, and motility using a “Ballistographic Chair.” These results suggest that cancellation tasks, such as those employed in the present comparative study, should also demonstrate age-related trends.

The purpose of these studies was to develop a set of normative data for a population of normal children, ages 6-14, on a simple letter cancellation task, and to compare these data with the performance of an ADHD population.

**METHOD**

*Materials*

The Mesulam Continuous Performance Test consists of two pages with the letters of the alphabet printed in uppercase. On one page (Ordered), the letters are placed in neat, orderly rows and columns, while on the second page (Random), the letters are placed in a haphazard fashion, with no apparent order imposed. On both pages, 60 A’s are placed among the other letters. Regardless of the page, the A’s are in the same location, dispersed symmetrically, with approximately 15 in each of the 4 quadrants. Figure I shows a portion of the Ordered and Random page.

Figure I
Study 1

Normative Sample

Normative data for the Mesulam were initially gathered from 1540 children (825 male, 715 female; age range 6 to 14 years, grade range 1 to 8). Participants were drawn from eleven schools (nine public and two parochial) located in three predominantly white, middle-class, urban and rural school districts in southern New Hampshire. The classroom teachers were surveyed to identify those children already diagnosed as, or suspected of, having learning difficulties. Approximately 11% (169 children) of the total sample were thus flagged by their classroom teachers as either having a specified learning disability, having a physical or cognitive impairment that prevented them from attempting the Mesulam test, or were formally identified as having an attention-deficit disorder with or without hyperactivity. These children were administered the Mesulam but were later excluded from the "normative" group, which resulted in a total of 1371 children included in this normative study. Testing was done at the end of one school year (May) and at the beginning of the next school year (September, 1993).

Procedure

All children were tested as a group in their classrooms. Separate grade level classrooms were chosen at random from the participating schools. Each child was given a crayon and the Mesulam protocol face down on the classroom desk. The children were given complete instructions that included the explanation that they were to "find all the ‘A’s and circle them. Do this as quickly as you can and when you think you have circled all that there are, turn your paper over." When the tester told the children to start, they were to pick up their crayon, turn the paper over, and begin the task. Because the children were tested as a group and not individually, and because this task was not considered to be a speeded task, the administration procedure did include the approximate but not precise timing of each completion. Children in all grades were limited to 7 minutes to complete the task (very few children went to this time extreme, and those that did were generally finished and simply rechecking the protocols at the end of the 7 minutes). The Mesulam was administered in a balanced manner, with 849 children being administered the Ordered page first (OP1 Administration), while 691 were administered the Random page first (RP1 Administration).

Measures of performance on the Mesulam include:

- **Ordered Page Errors (OPE):** This is the total number of "A"s that were missed on the Ordered page, regardless of administration order.
- **Random Page Errors (RPE):** This is the total number of "A"s that were missed on the Random page, regardless of administration order.
- **Total Errors (TE):** This is the total number of "A"s that were missed on both the Ordered and Random pages, regardless of administration order.

Each child’s protocol was scored for errors (missing an "A"). Ordered page errors (OPE), Random page errors (RPE) and Total errors (TE) were calculated for every student.
Obtained Score Differences on the Mesulam Continuous Performance Test

RESULTS

Means were computed for each of the nine age groups (6 to 14), and for both administration procedures (OP1 and RP1). The effects of sex, age, and order of administration were analyzed. No significant differences were found between groups according to sex. In contrast, highly significant effects were found for both page administration order and age. The effects of administration order and age were assessed by a two factor analysis of variance. OPE were found to be significantly affected by both administration order (F=6.709, P=.0097) and age level (F=14.654, P=.0001). RPE and TE were found to be non-significantly affected by administration order (RPE: F=.098, P=.7652, TE: F=2.934, P=.0869), but significant effects were noted for age level (RPE: F=10.379, P=.0001, TE: F=18.045, P=.0001). Because of these effects, separate normative tables based upon order of administration and age are provided. Table II presents data for the entire sample separated by age and order of page administration.

Table II: Mesulam CPT Means and Standard Deviations for Normative Population by Page Administration Order

<table>
<thead>
<tr>
<th>Age</th>
<th>n</th>
<th>OE M</th>
<th>SD</th>
<th>RE M</th>
<th>SD</th>
<th>TE M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>six</td>
<td>80</td>
<td>3.60</td>
<td>4.07</td>
<td>1.79</td>
<td>1.97</td>
<td>5.39</td>
<td>4.81</td>
</tr>
<tr>
<td>seven</td>
<td>124</td>
<td>2.90</td>
<td>3.23</td>
<td>2.30</td>
<td>2.96</td>
<td>5.20</td>
<td>4.93</td>
</tr>
<tr>
<td>eight</td>
<td>142</td>
<td>2.51</td>
<td>3.18</td>
<td>1.56</td>
<td>2.06</td>
<td>4.07</td>
<td>4.14</td>
</tr>
<tr>
<td>nine</td>
<td>102</td>
<td>1.40</td>
<td>1.62</td>
<td>1.32</td>
<td>1.63</td>
<td>2.72</td>
<td>2.40</td>
</tr>
<tr>
<td>ten</td>
<td>99</td>
<td>.91</td>
<td>1.21</td>
<td>.80</td>
<td>1.05</td>
<td>1.71</td>
<td>1.58</td>
</tr>
<tr>
<td>eleven</td>
<td>74</td>
<td>1.00</td>
<td>1.45</td>
<td>.77</td>
<td>1.35</td>
<td>1.77</td>
<td>1.94</td>
</tr>
<tr>
<td>twelve</td>
<td>76</td>
<td>.61</td>
<td>.93</td>
<td>.43</td>
<td>.79</td>
<td>1.04</td>
<td>1.19</td>
</tr>
<tr>
<td>thirteen</td>
<td>102</td>
<td>.25</td>
<td>.48</td>
<td>.31</td>
<td>.69</td>
<td>.56</td>
<td>.84</td>
</tr>
<tr>
<td>fourteen</td>
<td>50</td>
<td>.52</td>
<td>.89</td>
<td>.40</td>
<td>.76</td>
<td>.92</td>
<td>1.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>n</th>
<th>RE M</th>
<th>SD</th>
<th>OE M</th>
<th>SD</th>
<th>TE M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>six</td>
<td>23</td>
<td>1.96</td>
<td>1.85</td>
<td>.74</td>
<td>1.05</td>
<td>2.70</td>
<td>2.16</td>
</tr>
<tr>
<td>seven</td>
<td>56</td>
<td>1.98</td>
<td>2.28</td>
<td>1.89</td>
<td>2.05</td>
<td>3.88</td>
<td>3.04</td>
</tr>
<tr>
<td>eight</td>
<td>167</td>
<td>1.59</td>
<td>1.85</td>
<td>1.20</td>
<td>1.54</td>
<td>2.79</td>
<td>2.50</td>
</tr>
<tr>
<td>nine</td>
<td>71</td>
<td>1.49</td>
<td>1.71</td>
<td>1.13</td>
<td>1.34</td>
<td>2.62</td>
<td>2.26</td>
</tr>
<tr>
<td>ten</td>
<td>45</td>
<td>1.22</td>
<td>1.59</td>
<td>.98</td>
<td>1.42</td>
<td>2.20</td>
<td>2.19</td>
</tr>
<tr>
<td>eleven</td>
<td>51</td>
<td>.55</td>
<td>.92</td>
<td>.35</td>
<td>.52</td>
<td>.90</td>
<td>1.22</td>
</tr>
<tr>
<td>twelve</td>
<td>59</td>
<td>.54</td>
<td>.92</td>
<td>.17</td>
<td>.38</td>
<td>.71</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Errors on each page, regardless of administration order, demonstrated an expected age effect, with each variable generally improving from one age to the next. For all measures, scores are highest for the 6 to 9-year-old groups, and tend to level out through most of the other years. Order of page administration most highly affected the results for the younger children, ages 6, 7, and 8.

For ages 6, 7, and 8 there was a pronounced difference in the total number of errors (TE) made depending on the order in which the pages were administered. For example, six-year-olds given OP1 Administration had 5.4 total errors while those six-year-olds given RP1 Administration had an average of only 2.7 total errors. Interestingly, most errors occurred on the first page presented (Ordered or Random). Whichever page was administered to the child first yielded comparatively more errors than the succeeding page. Six-year-olds given OP1 Administration averaged 3.6 errors on the Ordered page, compared to the 1.8 errors on the Random page. Similarly, six-year-olds given RP1 Administration averaged 1.9 errors on the Random page, while averaging .74 errors on the Ordered page.

Because of the effects of age and order of administration on test performance, age-based norms for OP1 administration were selected for a sampling of an ADHD population.

### Study 2

**ADHD Sample**

170 children, between the ages of 6 to 14, were identified as having an attention-deficit disorder with or without hyperactivity. The children had been referred to an independent center connected to a major medical facility for evaluation of attentional disorders. The children were diagnosed as having ADHD on the basis of history, psychological interview by one of three psychologists, psychological/educational testing, and behavior rating scales. Determination of ADHD followed criteria set forth in DSM-III/IV. Included in the testing battery, but not used as a criterion for diagnosis, was the Mesulam CPT. All children were administered the Mesulam with the ordered page first (OP1). Table III shows the Ordered Page Error (OPE), Random Page Error (RPE), and Total Error (TE) means and standard deviations for the ADHD group. Also included are the results of two-tailed independent t-tests conducted to determine if there were differences in TE at each age between the control and ADHD groups.

#### Table III  AD/HD Sample Means and Standard Deviations for OP1 Administration and Two Tailed Independent t-Test Comparisons for Total Error - AD/HD Means vs. Normative Means

<table>
<thead>
<tr>
<th>Age</th>
<th>n</th>
<th>OE M</th>
<th>SD</th>
<th>RE M</th>
<th>SD</th>
<th>TE M</th>
<th>SD</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>seven</td>
<td>27</td>
<td>10.56</td>
<td>9.81</td>
<td>13.46</td>
<td>8.97</td>
<td>24.23</td>
<td>16.70</td>
<td>5.81 **</td>
</tr>
<tr>
<td>eight</td>
<td>32</td>
<td>6.88</td>
<td>7.50</td>
<td>9.68</td>
<td>7.57</td>
<td>16.52</td>
<td>13.03</td>
<td>5.32 **</td>
</tr>
<tr>
<td>nine</td>
<td>21</td>
<td>6.62</td>
<td>7.77</td>
<td>10.57</td>
<td>9.33</td>
<td>17.19</td>
<td>15.68</td>
<td>4.23 **</td>
</tr>
<tr>
<td>ten</td>
<td>32</td>
<td>3.97</td>
<td>2.98</td>
<td>6.16</td>
<td>5.74</td>
<td>10.13</td>
<td>7.24</td>
<td>6.58 **</td>
</tr>
</tbody>
</table>
The ADHD group as a whole had significantly more errors on each of the three measures than the normative group at each age level. Comparing the ADHD group’s performance on the Ordered and Random pages to that of the normative group showed an unexpected and reverse demonstration of errors. While the normative group typically did better on the second, Random page when given Ordered Page First (OP1) Administration, the ADHD group did not show this positive trend. ADHD children given OP1 Administration typically had slightly more errors on the second, Random page. At all ages except one (age eleven), the ADHD group committed significantly more TE than did the normative group.

**DISCUSSION**

This paper presents the results of norming procedures for the Mesulam Continuous Performance Test for a large sample of school-aged children. Results demonstrate the need for caution when administering such tasks, since minor changes (page administration order) had major effects on the results. This set of studies suggests that the Mesulam may potentially serve as a quick, inexpensive evaluative tool in the diagnosis of ADHD.

Evaluators in schools may find the Mesulam a useful tool in screening large numbers of children quickly and inexpensively. Both pages of the Mesulam can be easily group-administered to whole classrooms in less than 15 minutes. The normative data imply developmental trends in attentional capacity, thereby lending credence to its validity.

The Mesulam task should be considered only one component of the complete multi-method approach that is needed in making accurate diagnoses of this disorder.

Another possible limiting factor of this study is that this sample may not be representative of the broader population and additional norming may be needed. Evaluators may wish to develop their own "local norms".

The ability of the Mesulam to identify ADHD children from those with other disorders has not yet been established. Children may often display characteristics that resemble ADHD, but suffer from symptoms that are more appropriately attributed to other disorders. Multi-method evaluations should address whether the learning problem is a manifestation of a specific learning style, underlying process disorder, neurological deficit, or attentional disorder.

**REFERENCES**


Obtained Score Differences on the Mesulam Continuous Performance Test


Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Test behaviors of interest & Testing the Limits Ideas

John Willis, Jonas Taub, Ron Dumont and Nancy Marron

Examples are gleaned from the WISC-III and DAS but may be relevant to other tests that utilize similar measures:


**Information:** Is the student able to retrieve information fairly easily, or is there inaccuracy and inconsistency? Does the student have particular holes in his or her knowledge base? When they do, what about the difficult items they are successful with? Are there particular strengths or areas of keen interest and experience? Interview will reveal this. I am always observant of the kid who misses questions involving time concepts but has all this science knowledge.

Group items by category (Number items, science items, general factual items). Does the child miss a certain type?

**Testing limits:** With some children, who miss number items, I ask "Tell me the days of the week?" Can they names the days but out of sequence? Do they have a problem initiating? (Long pause with no answer. You say, "Let me start you - Sunday...." The child jumps in and says the rest). "Tell me the months of the year."

**Similarities (DAS - Similarities):** The DAS format is like the WISC-III with the exception that it provides 3 target words as opposed to 2. This allows a child who doesn't know one of the words to still have a chance of getting passing credit. How often do you get asked to define Tribe?

**Arithmetic:** Do they ask for repetition? Are they able to hold and manipulate information without losing some of it, especially Arithmetic and Digits Reversed. Do they remember the information correctly but get the task wrong, or is their memory inaccurate?

**Testing the limits:** try a paper and pencil task. First try presenting the same problems in written form. Then allow scratch paper. Finally, by converting the word problems to a paper and pencil calculation test (see example) you might be able to discern if this is a problem of attention/concentration and/or mathematical skill.
**Digit Span (DAS Recall of Digits):** Obviously make a verbatim copy of the response, whether correct or not. This allows for item analysis. What errors were made? Did the child simply forget the items, remember the items but mis-sequence them, or remember the items but have intrusions of other numbers. Did the child on the WISC-III verbally compensate by repeating the numbers to him/herself while you read them at a rate of 1-per-second? On the DAS the numbers are read 2-per second, somewhat alleviating the verbal compensation and making the test a "purer" measure of short-term memory.

**Picture Completion:** Does the child respond with a word, a description of the missing detail or by pointing? When he has difficulty coming up with the specific labeling word, I begin to wonder about word retrieval difficulties.

I use the abbreviations PC and PIC for "Pointed correctly" and "Pointed Incorrectly".

**Testing the limits:** I have used the PC subtest as an un-normed naming vocabulary test. Go back and instead of “What is missing” try “What is this”. If I sense poor one word expressive, I follow up with a better measure.

**Coding:** I observe their tracking skills and the errors they make. Do they have an orderly approach, or do they have difficulty finding their place?. Do they begin to memorize symbols (you'll notice this if they don't look up at the code on top)?

Beware of making much about speed and slowing. Some (Kaufman (1995), Nicholson, & Alcott (1994)) have suggested that slowing of responses on the Coding subtest may be indicative of certain problems. We [John Willis and Ron Dumont (1998)] found that slowing on Coding was normal.

**Picture Arrangement:** Time the standard procedure but also record the time it takes to initiate action. Although the student receives credit based on the total time for completion, many children spend large amounts of time "figuring out" the stories. Timing how long it takes to make the very first move may help understand how they are processing the information.

**Testing the limits:** Any story that is incorrect should be questioned as part of the testing-the-limits stage. I lay out the pictures exactly as the child arranged them and ask: "Tell me the story that this represents." I get a sense of the logic behind the arrangement.

**Block Design (DAS - Pattern Construction):** I am intrigued by students who misplace one of the diagonal blocks on Block Design. Seeing the error, the student rotates the block -- still wrong. A second rotation -- wrong again. The next turn will do it, but instead the student flips the block over to the identical, opposite face and starts over. I mark this behavior on the protocol as TTF (turn, turn, flip). Some kids rate several TTFs. I then wonder about frustration tolerance, ability to stick with a strategy, etc. As with all such hypotheses, I then seek real-life evidence to confirm or refute my guess. Sometimes it is a unique response to the artificial demands of the test. Occasionally it helps me learn something about the student's operating style in other situations.

I use CBO and CBR as abbreviations to indicate "Correct but overtime" and "Correct but rotated."

Look to see if there were quantitative as well as qualitative differences as the task changes. Note that Items 1-2 have the model made by the examiner (3 dimensional), items 3-5 have a model with the dividing lines shown (2 dimensional), items 6-9 have 4 blocks but no dividing lines in the picture, and items 10-12 have 9 blocks but no dividing lines. Some children will make errors as the stimuli shift slightly. I look to see if that is where the errors begin.

**General -** What is the student's response style? Quick or slow. Self confident. Impulsive. Cautious. Is there a response delay and what might it be due to?

When the student makes an obvious error -- e.g., an obviously senseless PA story; or a total mess of a BD item; or making a giraffe with a goiter instead a horse; or a VW with two extra, floating parts instead of a large sedan; or a Picasso face -- does the student make a visible effort to ignore the error (hoping perhaps I will, too?), try to rationalize it ("That's how I make horses!" "It's a little car"), try again, or frankly admit failure and give up?
Measures latency before beginning a response and silences during a response by discretely marking one dot each second while he waits [yes, he also uses his stopwatch to time the item correctly]. The protocol gives a visible record of the student’s speed, fluency, and spontaneity and allows you to determine if overtime responses on, for instance, PC went overtime with the initial delay or subsequent pauses.

**Verbal Expression Tasks** - Vocabulary, Similarities, Comprehension - I look for word retrieval, precision, the categorical term vs. explanation. Response delays. Organization and expression. Focused responses vs. circuitous responses vs. successive approximation of the correct response -adding information until they have answered the question.

If the child asks for repetition of an item or a word, if the task is not a timed task, ask "What did you think I asked." This may help distinguish between auditory misperception, confusion, and/or memory issues.

**Visual Construction Tasks** - Picture Arrangement, Block Design, Object Assembly - I note problem solving approach. Are they intuitive, perceptive, able to anticipate connections, or do they try piece after piece until they find connections that work? There are some kids who will not try anything until they are pretty sure that it will work, and others who will try endlessly illogical connections. How well do they recognize errors and how well do they use this information to make corrections? Do they have difficulty seeing the whole and reproducing the patterns, or do they have difficulty finding the correct orientation of individual parts, even when they clearly know what the solution should look like. I always allow the child to continue beyond time limits until they finish (correct or not) or give up. I score strictly by the time limits, but I want to know if they can solve it correctly, and how many they are able to solve correctly beyond time limits. If I have exceeded the discontinue criterion, I stop when it is clear the child is beyond their ability level. But I’ve frequently gotten kids who reach discontinue criterion early on, but continue to solve most or all of the remaining items, often within time limits. When testing the limits, see if the child can make a previously impossible, 9-block construction inside the box.

On **Achievement Tests**, I try to analyze errors to determine which skills are strong, which are missing, what trips them up, and what keeps them from progressing.

On **Reading Tests** - I want to observe the strategies the student uses for decoding and comprehension. I take detailed notes of misreadings and miscues so I can analyze strategies used. I look for phonological confusion, phonetic skills and holes, the ability to segment and blend, knowledge of sight words, and the ability to apply these skills to reading, both in isolation and in context. I also observe fluency and automaticity.

On Math Tests – I later readminister items that were missed because of misreading operations signs or making simple factual errors. I give the student a calculator and red pen and allow the student to correct missed and skipped items (but I try not to let the student know I will be doing this so I can avoid an incentive to skip difficult items). With younger students, I make or have them make a number line before readministering missed items.

On **Writing Tests** – I also collect a writing sample done with a word processor. If you use a test with two forms (e.g., PIAT-R NU, WIAT, TOWL-3), you can even compare and contrast the two efforts. I run the student's writing (both handwritten and word-processed, but not formal spelling tests) through one or more spelling check programs or machines. It is helpful to know what percent of the student's errors are picked up, how many of those elicit the intended word from the machine, and how many of those are given as the first choice. Térèse Murphy, Jaffrey-Rindge, NH, School District, found that different spelling check programs and machines had very different degrees of effectiveness for different individuals.

In verbal tests, does the student have word finding difficulties--does he "talk around" the answer? Does he have trouble phrasing sentences in correct grammatical order? Does she say she knows the answer but can't remember it right now? Does she give up too quickly, does she refuse to give up even if you have offered to go on to the next item? How does he react to not knowing an answer? On performance items, does he have an organized trial and error method or a more random style? On block design does he just place the blocks without trial and error? Can he rotate a row of blocks at the bottom when he sees they are the wrong direction or does he start over? Does he continue to repeat the same mistakes over and over? Has he learned anything from the successful completion of earlier test items such as in block design? Does he take apart puzzles or block designs that are correct? In spite of saying, "It's a soccer ball," does he continue to attempt to assemble the puzzle into "unround" shapes? Does he use only one hand or worse, use his right hand for the right side of the design and his left for the left side of the design? Does she show anxiety when she detects that speed is important? Does he make self-degrading remarks about his performance or say "this is so easy" when he is actually doing
Test behaviors of interest

very poorly? Does he make unusual movements like tremors, foot tapping, finger tapping, tics, hair twisting, or make repetitive noises, throat clearing, whispering to self, repeating phrases over and over?


Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.
Not All Test Scores are Alike

The Gift of Time

- Grade Norms vs. Age Norms (see Woodcock-Johnson chart)

Grade-Equivalent Scores

- Not equal units: cannot be added, subtracted, multiplied, divided, nor averaged.
- Do not reflect the student's actual functioning level (see OAT-CEREAL)
- May not be real scores at all (interpolated and extrapolated)
- May not even be grade levels included in the test.

Percentile Ranks

- The percent of students whose scores were tied or beaten by this student. The 37\textsuperscript{th} percentile rank means you scored as high as or higher than 37 percent of the students in the test's norming sample or in your local group. The 99\textsuperscript{th} percentile means you were in the highest one percent of the group.

- Nothing to do with percent correct. (Never use % sign in an abbreviation!)
- Not equal units: cannot be added, subtracted, multiplied, divided, nor averaged.

Standard Scores and Scaled Scores

- Measure how far the student scored from the average in terms of the average spread of scores for the whole group. A standard score of 115 or scaled score of 13 means the student scored one standard deviation above the average (which would be the 84\textsuperscript{th} percentile rank). A standard score of 85 or scaled score of 7 means the student scored one standard deviation below the average (which would be a percentile rank of 16).

- Equal units: can be added, subtracted, multiplied, divided, or averaged if you're in the mood.
- Too narrow: Encourage obsessive comparisons between essentially identical scores.
- Often misunderstood.
Stanines

- Almost equal units: can be added, subtracted, multiplied, divided, or averaged if you wish.
- Too broad: Encourage obsessive comparisons between essentially identical scores.
- Fairly easy to explain and understand (on a good day).

Percentile Ranks and Standard Scores

- These two statistics will not always tell the same story. A student may not be many points away from the average and still have an extreme percentile rank or may be many points away from the average and nonetheless have a fairly average percentile rank.

Confidence Bands

- Test scores are never perfectly accurate. Lucky or unlucky guesses, lapses of attention, and other factors mean that the same person would almost never get exactly the same score on a test twice in a row. A confidence band around a score tells how scores on that test are likely to vary by pure chance.
- If the confidence bands on two scores overlap, there probably is not a significant difference between the two scores. On another day the higher and lower scores might have been reversed.
- If the confidence bands on two scores do not overlap, and if both scores are probably valid, there probably is a significant difference between the two scores. On another day, the higher and lower scores would probably have still been the higher and lower scores, respectively.

In the example above, there is a triumph of Hope over Experience, but neither is significantly different from Dumb Luck.

Stanines

- As a rough approximation, you can usually assume that two valid scores that are in adjacent stanines may not be significantly different, but that two valid scores that differ by more than one stanine probably are significantly different.

Significant Difference
A "significant difference" is one that is too large to have been likely to have occurred by chance when there was no real difference between the abilities being tested. This likelihood is expressed as a probability. e.g., p<.05 means that there were fewer than 5 chances in 100 of a difference that large or larger happening by accident.

**Base-rate**

- Base-rate refers to the prevalence or frequency of a particular occurrence or event within a population. Awareness of relevant base-rate data allows an evaluator to determine the diagnostic utility of a particular sign. Although a particular relevant comparison may reach some level of statistical significance, it is always necessary to determine if the statistical difference is a usual or an unusual one. Base-rate information provides just such data.
- Although an 11.2 point difference between scores on the WISC-III Verbal and Performance scales represents a statistical difference, base rate tells us that such an occurrence is likely happen in about 40.2% of the population.

**Item Analysis**

- A score can tell you only so much, and some of what it tells you may be wrong. To really understand a student's test performance, you need to look at the individual item responses. For example, the Gates-MacGinitie manuals have extremely valuable sections on sources of error, such as responding to a single word in the paragraph instead of the whole text, making false assumptions on the basis of prior knowledge, or overemphasizing one part of the story.

**Scoring**

- All scores should be done three times. Count the number right. Then count the number wrong and subtract that from the total. Finally, start with the number right and add one point for each wrong item; you should come out with the total number of items at the end. Read numbers and titles of tables, columns, and rows aloud as you look up scores and listen to yourself.
- Examine your resulting standard scores. Do any look like they don't belong. The child with a score of 143 on Memory for Sentences and scores hovering around 100 all on other test should set off a mad rush to answer the question "Why?" Often you will find that you made a mistake when you scored the test as opposed to having just discovered some weird ability of the child. "Tester: Blame thyself before passing the blame on to others."

**TEST SCORES ARE NOT NECESSARILY TRUSTWORTHY**

**PERVERSIVE INVALIDITY**

- Princess Summerfallwinterspring and seasonal norms.
- The student may have just blown off the test.
- The student may have had a bad day.
The student may have followed the instruction to skip too-difficult items but forgotten to skip the corresponding items on the answer sheet.

The student may not have followed the instruction to skip too-difficult items and spent most of the time struggling bravely but fruitlessly on one impossible item.

The answer sheet may have baffled the student (hint: if the name is spelled wrong on the print-out, the scores may well be invalid).

The student may fail to switch tasks (e.g., initial sounds to final sounds, synonyms to antonyms).

The student may be carrying out an entirely different task from the one intended.

The ordinarily very generous time limits may be too short for a few students who work very slowly.

The ordinarily reasonable time allotments for subtests may exceed some students' attention spans.

The students score may be low but it results from doing what has been instructed in the classroom. The student "worked slowly" but accurately thus completing very few Coding (Symbol Search, Cross Out, etc.) items but getting every one correct.

The students score may be low but it results from doing what has been instructed on the test. The student "worked quickly" in but accurately thus completing very many Coding (Symbol Search, Cross Out, etc.) items but getting many incorrect.

CONFUSION BETWEEN INCAPACITY AND SPECIFIC PROBLEMS

Free-response and multiple-choice tests are not comparable for some students.

The student may have misread operation signs.

The student may know the process (e.g., long division) well but make computational errors (e.g., subtracting wrong in an otherwise correct long division problem).

The student may fail otherwise easy math applications problems because of reading difficulty.

The student may understand fairly high-level skills but make simple errors on much simpler skills (see the OAT-CEREAL).

The score may slightly overestimate the student's working level if the student is unusually accurate on the problems the student can solve.

Study and use the interpretive suggestions in the test manuals.

Content on these pages is copyrighted by Dumont/Willis © (2001) unless otherwise noted.