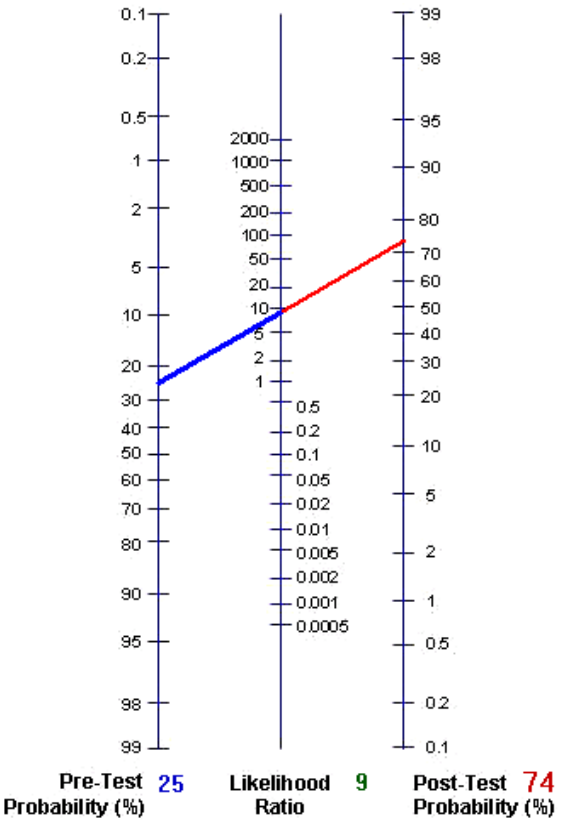
Implementing Evidence-Based Practices in the Assessment of ADHD

Thomas K. Pedigo, ED.D.
Vann B. Scott Jr., Ph.D
Stephen J. Hughes, Ph.D.



n =555	ADHD	TYP	Totals
Test Positive	265	34	299
Test Negative	26	230	256
Totals	291	264	555

Poster presented at the 5th Annual Conference
American Academy of Clinical Neuropsychology June 7-9, 2007 Denver,CO

Targeted Testing Inc. ™
www.targettest.com

Abstract

The American Academy of Pediatrics (AAP) and the National Institute of Health (NIH) have both rendered statements on the appropriate assessment and treatment of Attention-Deficit Hyperactivity Disorder (ADHD; AAP, 200X, NIH, 200X). Both organizations called for the creation of assessment tools with strong clinical and psychometric properties for use in primary care and traditional assessment settings. This poster describes the Pediatric Attention Disorders Diagnostic Screener (PADDS), an ADHD assessment suite developed in response to this call. The PADDS consists of a semi-structured diagnostic interview and three computer-administered measures of executive functioning, the Target Tests of Executive Functioning (TEF). The system also features a dynamic, computer-generated nomographic display that illustrates the incremental power obtained through combined use of the TEF tasks. Remarkable diagnostic accuracy is obtained through use of the TEF in an evidence-based assessment model (Frazier, 2004, 2006). For example, when screening cases from a population with an assumed base rate of 4% for ADHD, the PADDS system can produce an accurate diagnostic hit rate of as much as **99%.** The PADDS system shows clear diagnostic superiority over conventional standard score or reference-group-comparison approaches, and addresses the need for strong, empirical assessment tools expressed by the AAP and NIH.

Incremental Validity / Evidence-Based Assessment

Development of the PADDS was undertaken through the approach described by Hunsley and Meyer (2003) and Johnston and Murray (2003) of demonstrating *incremental validity*. Our presentation of the evidence we have gathered begins by directly addressing a series of issues inherent in the incremental validation of instruments used to diagnose children and adolescents outlined by Johnston and Murray (2003) including: a) developmental and contextual issues, b) problems with defining criterion variables, c) determining base rates, and d) specificity issues involved in establishment of incremental validity, particularly in the case of a newly developed instrument.

Evidence of the Incremental Validity of the PADDS in the diagnosis of ADHD

➢**Developmental and contextual issues**. Johnston and Murray (2003) note that when developing of a new diagnostic instrument, researchers must take into consideration the developmental status of the individuals for whom the measure is intended as well as review the incremental validity that the diagnostic tool provides across the developmental stages of the individuals for whom the measure is intended. While developing the PADDS, we carefully examined the utility of the system as a function of the age of the children tested, and determined that the system loses utility as the age of the child being tested moves beyond 12 years – a function of the difficulty level of the tasks and the developmental status of individuals at this age and beyond. We therefore determined that the incremental validity in using the PADDS diminishes after age 12.

➢ **Defining criterion variables**. As noted by Johnston and Murray (2003), "Demonstrating incremental validity will always be easier when the criterion is objective, widely agreed on, and not confounded with the information used in prediction." (p. 499). The diagnostic status of the known-ADHD, and unaffected control groups was carefully established independent of performance on the PADDS through comprehensive psychological evaluation and clinical interview as well as the use of widely used measures commonly used to assess symptoms of ADHD including various parent and teacher ratings scales of DSM-IV criteria for ADHD, such as Connors Rating Scale or Brown ADD Scales. Typicals and non-typicals were screened using the Connors Rating Scale (using the parent and/or teacher versions). Students accepted for inclusion as typical were those with T-score rating below 65.

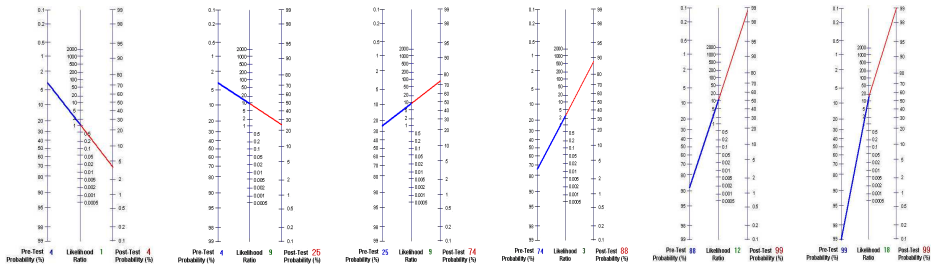
➢**Determining base rates.** A particularly challenging task involved in the establishment of incremental validity is determining the appropriate base rate against which to make the comparison. Indeed, the results of a positive test are far less surprising if the client is drawn from a population with a high incidence of the disorder. Therefore it is imperative to incorporate base rate information into the process of establishing incremental validity. In our work with the PADDS, we adopted the most current estimates of base rate incidence of ADHD as the default value for the base rate calculation in our analyses. In addition, we have, as will be explained in more detail later, established a mechanism by which base rate data for an individual site's or practitioner's referral base can be incorporated into the overall assessment process.

➢**Specificity of incremental validity**. Incremental validity is highly dependent upon the context in which source data are collected and applied. We adopted an evidence-based procedure described by Frazier & Youngstrom (2006) to utilize **likelihood ratios within a nomogram** to improve the appropriateness of decisions based on the multiple sources of data. Frazier & Youngstrom (2006) recommended that **nomograms** be utilized so that each data source can be allowed to contribute to (or detract from, as the case may be) the prediction of the diagnosis.

Likelihood Ratios and the Nomogram

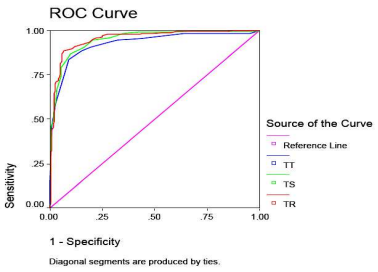
The likelihood ratio, (LR), is a key concept in the interpretation of diagnostic test results. A LR **combines** the values of ***sensitivity* and *specificity*** into a *more* ***meaningful number*** and allows direct estimate of how a test result changes the odds of a positive or negative diagnosis. The LR is used to assess how good a diagnostic test is and to help in selecting an appropriate diagnostic test or sequence of tests. LR's have real advantages over both sensitivity and specificity alone because they are less likely to change with the prevalence of a disorder, they can be calculated for several levels of a test result, and most importantly, ***they can be used to combine results of multiple tests***, and to calculate a post-test odds, or probability for a target disorder based on the individual's test performance.

The examples below show a stepwise calculation of post-test probabilities derived using likelihood ratios, using data from the case study, applied incrementally on Fagan's Nomogram. Beginning with a conservative base rate of ADHD prevalence at 4%, It clearly illustrates the power and utility of this approach.



METHODS

PADDS* *Target Tests of Executive Functioning (TTEF) are computer based screening measures of attention and executive functions with proven clinical reliability and validity for the screening of attention and executive disorders (Pedigo & Scott 2007; Pedigo, 2005; Scott, 2005; Pedigo, 2003). As part of a multi state/multi site normative study, the TTEF were administered to 555 children ranging in age from 6 to 12, with a mean age of 8.58 years *(SD* = 1.67). Within this sampling, well-defined reference groups were established from sites in four states (Idaho, Georgia, Tennessee and California), including 299 ADHD children and 256 age matched non-ADHD children. The sample included 319 males and 236 females, and was composed of: 222 African Americans, 278 Caucasians and 55 Hispanics. ADHD children were removed from their medication for the morning of the TTEF subtests administration. The typical children were drawn from public and private schools and screened for ADHD with a Conner's teacher and or parent rating prior to inclusion and administration of the TTEF. All subjects received $10 as a gift for participation in the study. Analyses were undertaken to calculate ROC, sensitivity, specificity, and positive and negative predictive power. The ROC curve (seen at left) is a visual representation of the accuracy of the discriminating ability of the TTEF and the values for the area column (area under the curve) are to be interpreted as accuracy scores. Asymptotic sig. represents the probability that the resulting accuracy of a diagnostic test is only a chance occurrence (i.e., the values are to be treated as p values in a test of statistical significance).



Area Under the Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
TR	.954	.010	.000	.935	.974
TS	.948	.010	.000	.928	.968
TT	.927	.013	.000	.901	.952

a. Under the nonparametric assumption

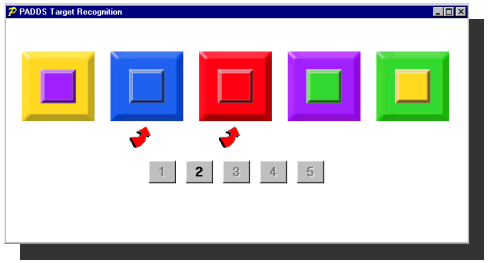
b. Null hypothesis: true area = 0.5

PADDS Psychometric Performance

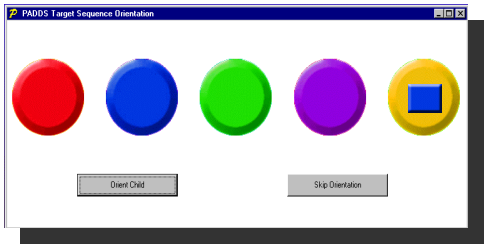
n =555	ADHD	TYP	Totals	
Test Positive	265	34	299	PPV =0.89
Test Negative	26	230	256	NPV =0.90
Totals	291	264	555	Sensitivity =0.91
				Specificity =0.87

PADDS Case Study / Discussion

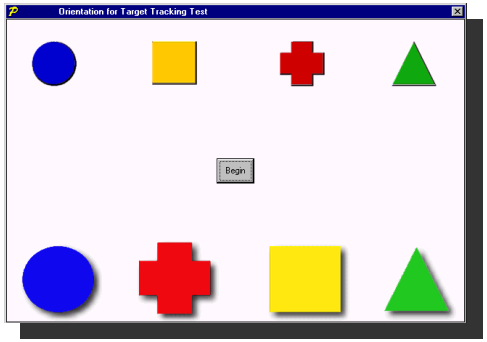
As the "TARGET TESTS OF EXECUTIVE FUNCTIONS" are relatively new, they are given additional description here:



Target Recognition- presents five large colored squares with smaller squares inside them. Below the squares are five buttons marked 1 thru 5. The colored squares simultaneously blink on and off at 1 ½ second intervals over 153 presentations. Each cycle shows a different combination of matching colors. The child is taught a strategy to count from left to right, the number of squares with small squares of the same color and to click on the corresponding numbered button.



Target Sequencing- Across 39 trials, Five colored circles have a small square appear and disappear in varying sequences. The child is taught to attend only to circles when the square matches it in color and to remember the sequence of color matches, first match first, second match second and last match last. Then to click on the circles that matched in the same order as they were presented.



Target Tracking - Across 20 trials this subtest presents four colored shapes at the top and bottom of the screen. The shapes will move, one at a time, to another shape at the bottom in differing sequences of two and three step moves, then reset. The child must then click and drag each shape to it's position on the lower shapes, in the same order. First move first, second move second and last move last.

PADDS Case Study / Discussion, continued...

Current Case Summary: TP is a 6 years 1 month old female that presented with a history of falling grades. She has demonstrated above average achievement potential on previous standardized testing and made honor roll grades in kindergarten. Since the Christmas break she has struggled to complete assignments, and has begun making careless errors. At present she is in jeopardy of repeating 1st grade. TP's mother reported that it often takes several hours to complete home work that should be done in far less time. There are problems with follow through of routine instructions as well. TP's mother reports having similar experiences as a child but was never formally evaluated. Her mother completed the Computer Assisted Diagnostic Interview (CADI) to assist with historical report. Development was described as typical. There were no major medical or psychosocial issues reported. TP demonstrated a normal range of age appropriate interests, is described as very social and well behaved. There were no reports of emotional disturbance or psychopathology. A review of DSM-IV diagnostic criteria for ADHD was completed by both mother and teacher and was suggestive of a formal attention disorder.

Techniques Administered: Parent and Child Clinical Interview; Review of school records, Parent and Teacher BASC-2, Parent and Teacher BRIEF, WASI, WRAML, PADDS, TOVA.

Summary of Results: WASI FSIQ- 113, WRAML-2 GMI-95, TOVA ADHD SCORE-4.60, Target Recognition-109, Target Sequence-17, Target Tracking-4, BASC-2 areas of agreed significance Attention problems T-score 68/74 Parent and teacher respectively , BRIEF areas of agreement-multiple areas of core Executive skills for both Parent and Teacher.

Target Tests of Executive Functioning: The computer tasks were designed to tap abilities often needed for successful performance in the routine classroom by requiring greater aspects of executive functioning, i.e. planning, organization, working memory, attention to detail/changes in stimuli as opposed to merely relying on boredom as in persistence tasks.

Results and Conclusion

Information from the above case was input with a conservative estimate of a base rate of 4%. A recent review of the DSM-IV Rating Scales estimated diagnostic accuracy for ADHD completed by Zolotor and Mayer (2004) demonstrated accuracy to be between 90 to 95 percent for both. However, ratings are not always in agreement and are known to be susceptible to bias. Thus, to remain consistent with our conservative model each rating was assigned a likelihood Ratio (9) consistent with a sensitivity and specificity of 90% respectively.

Results and Conclusion

As can be seen in our conservative model when applied against the base rate of 4 %, a positive parent rating improved our post test predictive index to 25% and when the teacher rating is added we achieve a post test predictive index of 74%. This is considered an important contrast to the predictive assumption (90-95% diagnostic accuracy) that is suggested when using overall sensitivity and specificity of measures that were obtained from studies assessing the ability of that measure to reassign subjects with known group classification (100% versus 4%). Thus, as shown in this example, the ability to obtain objective samples from multiple sources (TR, TS, TT) of information can clearly lead to improvement in the predictive model. Additionally, this process can serve to add incrementally validity to diagnosis via cross validation of the behaviors ratings. The difficulty has been finding such measures that also include well defined clinical and non-clinical subjects with comparable psychometric performance (sensitivity, specificity, positive and negative predictive power) from which Likelihood Ratios can be derived for every possible outcome. The Target Tests of Executive Functioning were designed to fill this need. For example, a review of the predictive power for just the Target Tests in this case (without the behavior ratings) produced an overall predictive index of 94%.

The overarching emphasis in this proposed approach is to use several measures with solid psychometric properties and proven clinical utility in the most conservative manner possible. The beginning premise of our case study assumes that prior to any input or analysis there is a 96% chance that no clinical disorder exists. This assumption is then analyzed in incremental fashion allowing for both additive and subtractive information to be collected in producing an overall predictive index. While overall Predictive Index in this case (99%) is impressive it is suggested that this information should then be applied along with other measures and clinical judgment of co-morbidity to made by the clinician in answering the following question: All information considered, what is the risk of rendering a diagnosis and recommending medical intervention versus doing nothing. As can be seen in this case there is clear academic underachievement that has likely been ongoing and increasing and there is a frequent parent and child flash point that occurs around the struggle with home work and the amount of time and supervision it takes to complete. In this instance it was felt that the combined evidence justified a diagnosis and ADHD Primarily Inattentive Type along with a recommendation of a trial of stimulant intervention. Follow up of this case revealed low dose long acting stimulant produced significant improvement for task completion and improving grades followed by a very low dose in the PM to assist with homework issues.