What is the science behind screening and how can we recognize well-constructed, accurate tools?

---

**“High-Quality Developmental Screening Instrument” Defined**

High-quality developmental screening tools are those that are standardized, reliable, valid, and accurate (that is, both sensitive and specific, correctly detecting children with and without problems).

<table>
<thead>
<tr>
<th>Definitions</th>
<th>How PENDS meets these*</th>
</tr>
</thead>
</table>
| For **standardization**, this means that the screening tool was standardized on a large nationally representative population (not a referred population) and with at least 100 subjects per age range | PEDS was initially standardized in both English and Spanish in 1997 on 771 children between 0 – 8 years of age from 5 states representing the broad geographic locations of the US. Since 1997, PEDS has been the focus of three additional standardization studies. One involved 408 children between 0 – 2 years of age in 22 states. This study included, for the first time, Native Americans and Pacific Islanders. A second study involved 525 children between 0 and 6 years of age in 17 different states. A third study included 1119 children between 5 and 8 years of age in five different states. Across all studies for each sample and for the total sample of 2823 children, families were representative of the US population in terms of ethnicity, parental levels of education, income, whether they lived in urban, rural, or suburban areas, frequency of developmental disabilities (only 26 out of 130 had been previously detected).

Across studies, a variety of sites were used including: pediatricians’ offices and medical centers, day care centers, schools, families participating from their own homes. At each site, consecutive patients, or all children in a classroom (and their unenrolled siblings), were used to ensure that sampling was realistic. A variety of methods were used to elicit parents’ responses to PEDS: telephone interviews, paper-pencil self-completion, and face-to face interviews, although in most cases, parents’ completed PEDS independently.

Additional studies are ongoing in both the US and elsewhere. Most focus on translation into other languages (now including Somali, Vietnamese, Hmung, and Chinese) and standardization of PEDS in other countries, notably Canada, Slovenia, Australia, Malaysia, etc. |
| For **reliability**, this means correlations of 0.85 or above for internal consistency, inter-rater consistency | As part of developing PEDS, three reliability studies were conducted. The first viewed, using the responses of 20 different parents, the extent to which their concerns could be reliably categorized by two independent observers. Agreement ranged from 80% to 100% across categories, with an average of 95%. This illustrates that the PEDS has a high degree of inter-examiner, also |
rater consistency, and test-retest reliability.

called inter-rater reliability.

A related question is whether two different interviewers could elicit from parents the same kinds of concerns. This was assessed by having a different examiner re-interview 40 parents and then comparing results from the two interviews. Although it would have been more desirable to re-interview parents on the same day, this assessment was conducted two weeks later and over the telephone instead of face-to-face. Nevertheless, agreement ranged from 80% to 100% and produced an average of 88%. This shows that PEDS can be reliably scored by different examiners.

Finally, coefficient alpha was produced to view the internal consistency of PEDS items and parents responses—as an indicator of homogeneity of content. The alpha produced on PEDS data was moderately high (a = .81). This suggests that parents’ responses to each item have reasonable levels of consistency. This also means that only a very small amount of variance in parents’ concerns is attributable to measurement error.

For validity, this means correlations of 0.70 or higher for concurrent validity and discriminate validity, and, if possible, evidence of predictive validity.

To assess various types of validity, each of the 771 children participating in PEDS validation studies was administered a battery of tests. For the majority of children, the battery was diagnostic and included measures of IQ, language, academics/preacademics, motor skills, social and emotional skills. The rationale for selecting a broad battery was to ensure that all aspects of development were measured.

To test concurrent validity (the relationship between parents’ concerns and subtests of the various measures, correlations were produced. These revealed a range of scores from .43 (between unlike content) to .91 for related content). For each type of concern, at least one concurrent measure resulted in a correlation above .75.

To assess discriminant validity, criteria were applied to children’s performance on the concurrent battery in order to discern the presence of various types of disabilities. The criteria were drawn from the US federal laws that ensure public school special education services for children with disabilities. Logistic regression revealed unique patterns of concerns were associated with various disabilities and that children with mental retardation, language-impairment, learning disabilities, physical impairments, autism and emotional disorders odds ratios and percentages were 8 to 13 times more likely than children without such disabilities to have parents with distinct patterns of concerns. In terms of decision-making based on this information, 79% of children needed comprehensive work-ups such as might be expected for those with mental retardation, learning disabilities or autism could be identified by the presence of certain concerns, while 75% of children needing
speech-language evaluations could be identified by other patterns of concerns.

Finally, PEDS was the focus of a predictive validity study conducted by researchers at the Royal Children’s Hospital in Melbourne, Australia. This revealed that parents with concerns known to be predictive of disabilities about their kindergarten age children, had children who two years later, had substantial difficulties in school. This suggests that PEDS results should be taken seriously and needed intervention sought promptly.

For **accuracy** (also known as “criterion-related validity”) this means:

<table>
<thead>
<tr>
<th>AGE</th>
<th>SENSITIVITY</th>
<th>SPECIFICITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>0 - 1</td>
<td>3/4</td>
<td>66/82</td>
</tr>
<tr>
<td>1 - 3</td>
<td>27/34</td>
<td>117/149</td>
</tr>
<tr>
<td>3 - 4</td>
<td>26/35</td>
<td>118/165</td>
</tr>
<tr>
<td>4 - 8</td>
<td>42/57</td>
<td>172/245</td>
</tr>
<tr>
<td>TOTAL</td>
<td>98/130</td>
<td>473/641</td>
</tr>
</tbody>
</table>

PEDS, like all other screens, do contain error but it is anticipated that with repeated screening (the wise recommendation of the American Academy of Pediatrics) that any children not detected by a screen at time 1 would be identified in a subsequent application. Over-referrals also occur but research on PEDS shows that these children tend to perform in the below average range on those measures that best predict school success: intelligence, language, and academics/preacademics. So identifying them is worthwhile because they can benefit from programs like Head Start, Early Head Start, quality day care, afterschool tutoring, summer school, etc.

The original research on PEDS involved four cross-validation studies each of which was published in peer-reviewed journals of pediatrics and early childhood special education. A number of subsequent studies in peer-reviewed journals have refined and expanded the decision-making properties of PEDS. Please see www.pedstest.com for a list of studies and in many cases links to the journals where they were published. PEDS research is ongoing and subsequent studies from other authors are also catalogued in the PEDS comprehensive manual, Collaborating with Parents available at www.pedstest.com.
accurately reported— including publication in a refereed professional journal.

<table>
<thead>
<tr>
<th>Definitions</th>
<th>How the M-CHAT (Modified Checklist for Autism in Toddlers) meets these</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>The M-CHAT is an extension of the CHAT which was developed on Great Britain on 16,000 children across several different studies. Consistently, certain items served in the identification of children with autism spectrum disorder while also correctly identifying children with other kinds of disabilities including language impairment and mental retardation. The M-CHAT includes 9 items from the CHAT plus 21 new items.</td>
</tr>
<tr>
<td>Standardization</td>
<td>1101 children participated in the norming studies for the M-CHAT and were consecutive patients receiving well-care at pediatric practices or referred for screening at early intervention services.</td>
</tr>
<tr>
<td>Reliability</td>
<td>Chronbach’s alpha was produces for all items and was found to be high (I = .85)</td>
</tr>
<tr>
<td>Validity</td>
<td>Concurrent studies deployed a comprehensive battery including the Bayley Scales of Infant Development, The Vineland Adaptive Behavior Scale, and the Childhood Autism Rating Scale. There were significant differences in the performance of children with autism spectrum disorder (ASD) on all items of the M-CHAT with the exception of an item tapping enjoying of rocking and swinging, and ability to walk.</td>
</tr>
<tr>
<td>Accuracy</td>
<td>The M-CHAT was sensitive to the presence of ASD at 87%, and specific to non ASD at 99%. Over-referrals, while minimal given that positive predictive value was 80%, tended to be children with other developmental disabilities.</td>
</tr>
</tbody>
</table>

References


Glascoe FP, MacLean WE: How Parents' Appraise Their Child's Development. Family Relations. 1990;39;280-283.


Glascoe FP. Parents' Concerns about Children's Development: Prescreening Technique or Screening Test? Pediatrics, 1997;99:522-528.


Glascoe FP. Toward a Model for An Evidenced-Based Approach to Developmental/Behavioral Surveillance, Promotion and Patient Education. Ambulatory Child Health, 1999, 5; 197-208

Glascoe, FP. The Validation and Standardization of Parents’ Evaluations of Developmental Status. Diagnostique, 1999;23: 185-203

Glascoe FP. The value of parents’ concerns in early detection of developmental and behavioral problems. Child: Health Care and Development 2000;26:137-149,


Glascoe FP. Teacher’s Global Ratings and Students’ Academic Achievement, Journal of Developmental and Behavioral Pediatrics, 2001;22;163-168


Silverstein M, Sand N, Glascoe FP, Gupta B, Tonniges T, O’conner K. Pediatrician Practices Regarding Referral to Early Intervention Services: Is
Having an Established Diagnosis Important?, *Pediatrics*, submitted


King T. Glascoe F. Developmental Surveillance of Infants and Young Children in Pediatric Primary Care, *Current Opinion in Pediatrics* 2003, 15:624–629


Glascoe FP, Developmental Screening. Website pages for the AAP Section on Developmental and Behavioral Pediatrics, 1998


Glascoe FP. Uso De La Evaluación Por Los Padres Del Estado De Desarrollo Para Detectar Y Tratar Problemas Del Desarrollo Y Del Comportamiento En Atención Primaria. MTA Pediatria, 2001


Glascoe FP. Kundell S. How to Improve Patient Flow, Satisfaction, and Quality of Care Patient Care, 2002;36:77-80, 83-4,


Glascoe FP. Developmental and Behavioral Screening. In JA Mulick, JW Jacobson, (eds). Handbook of Mental Retardation and Developmental Disabilities:

Glascoe FP. Developmental-Behavioral Screens for Primary Care. AAP News, January, 2003


For more details on PEDS research, please go to www.pedstest.com. There are links to various articles and information on ordering the comprehensive manual: Collaborating with Parents: Using PEDS to Detect and Address Developmental and Behavioral Problems.