Behavioral Testing With Children

Victoria Molfese
Dennis Molfese
Jennifer Beswick
Jill Jacobi-Vessels
Behavioral Testing Provides Valuable Information
Brain Functioning Explanations of Behavior is Tempting
### Names, Numbers, and Location of the Organs

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>No.</th>
<th>Name</th>
<th>No.</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amativeness</td>
<td>13</td>
<td>Self-Esteem</td>
<td>26</td>
<td>Size</td>
</tr>
<tr>
<td>2</td>
<td>Conjugal Love</td>
<td>14</td>
<td>Firmness</td>
<td>27</td>
<td>Weight</td>
</tr>
<tr>
<td>3</td>
<td>Parental Love</td>
<td>15</td>
<td>Conscientiousness</td>
<td>28</td>
<td>Color</td>
</tr>
<tr>
<td>4</td>
<td>Friendship</td>
<td>16</td>
<td>Hope</td>
<td>29</td>
<td>Order</td>
</tr>
<tr>
<td>5</td>
<td>Inhabitiveness</td>
<td>17</td>
<td>Spirituality</td>
<td>30</td>
<td>Calculation</td>
</tr>
<tr>
<td>6</td>
<td>Continuity</td>
<td>18</td>
<td>Veneration</td>
<td>31</td>
<td>Locality</td>
</tr>
<tr>
<td>7</td>
<td>E. Vitativeness</td>
<td>19</td>
<td>Benevolence</td>
<td>32</td>
<td>Eventuality</td>
</tr>
<tr>
<td>8</td>
<td>Combativeness</td>
<td>20</td>
<td>Constructiveness</td>
<td>33</td>
<td>Time</td>
</tr>
<tr>
<td>9</td>
<td>Destructiveness</td>
<td>21</td>
<td>Ideality</td>
<td>34</td>
<td>Tune</td>
</tr>
<tr>
<td>10</td>
<td>Alimentiveness</td>
<td>22</td>
<td>B. Sublimity</td>
<td>35</td>
<td>Language</td>
</tr>
<tr>
<td>11</td>
<td>Acquisitiveness</td>
<td>23</td>
<td>Mirth</td>
<td>36</td>
<td>Causality</td>
</tr>
<tr>
<td>12</td>
<td>Approbativeness</td>
<td>24</td>
<td>Individuality</td>
<td>37</td>
<td>Comparison</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>Form</td>
<td>C</td>
<td>Human Nature</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td>Suavity</td>
</tr>
</tbody>
</table>

History of Phrenology on the Web
But Research Using Brain and Behavior Measures Is Ideal!
The Link

- Interpreting the link between brain waves and brain processing is helped by behavioral testing.

- Moving from theory/speculation to hard data in the increasingly “evidence-based” world.
THANK YOU FOR BACKING UP YOUR OPINIONS WITH HARD DATA.
What Is “Behavioral” Testing?

Assessments of children’s abilities
- Intelligence - domains of abilities (verbal, nonverbal); standardized assessments
- Cognitive skills - specific aspects of abilities (language, reading, memory, mathematic) standardized or unstandardized
- Constructed (latent) variables and “batteries”
  - summary scores or factor scores
  - selected subtests from a variety of instruments
- Behavior-Based (observed) Assessments
Choosing Behavioral Tests

Choice depends on the purpose for which the information will be used
- What might the brain tell us about children’s behavior?
- What might behavior tell us about brain processing?

This is Neuropsychology!
Ideal Tests

In a perfect world

- Behavioral tests are linked to specific areas of brain processing identified by ERP, fMRI, MEG,
  - The “frontal lobe” executive functioning tasks - Adel Diamond
- Brain and behavior tests confirm a pathology or area of processing
- Test performance confirms or disconfirms whether this child, on this test, fits the brain model
This Isn’t A Perfect World!

Next best thing to a perfect world: Test a model to get converging data

- For example, we are seeking to link a model of brain processing of speech sounds with development of cognitive abilities that build on this processing

- Speech sounds ➔ language skills ➔ reading
The Model

- Phonological processing of consonant vowel speech and speech-analogue sounds

- ERP data showing areas of brain processing of speech sounds
  - previous research

- ERP data showing amplitude and latency effects are linked with different areas of brain processing

- ERP processing is linked to differences in children’s behavioral performance on language and reading tests
The Research Questions

Can ERP responses to speech sounds be linked to behavioral performance in infancy, preschool period (3-4 years) and at school age (5 years +)?

- Are there developmental changes?
- Are there differences due to: SES, activities in the home environment, attendance at preschool, health and sleep habits?
- Are there individual differences?
Many of the questions that researchers are seeking to answer require longitudinal designs
- developmental change (longitudinal design)
  versus
- developmental differences (cross sectional design)
Longitudinal Research

- One of our projects studied children from birth until age 13 years.
- Age span required a number of different tests and methodological adjustments to fit with the children’s ages.
- The challenge is to select measures that provide evidence of development.
# Behavioral Testing

<table>
<thead>
<tr>
<th>Test</th>
<th>Birth</th>
<th>1 Yr</th>
<th>2 Yrs</th>
<th>3 Yrs</th>
<th>4 Yrs</th>
<th>5 Yrs</th>
<th>6 Yrs</th>
<th>7 Yrs</th>
<th>8 Yrs</th>
<th>9 Yrs</th>
<th>10 Yrs</th>
<th>11 Yrs</th>
<th>12 Yrs</th>
<th>13 Yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERPs</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Bayley MDI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peabody Picture Vocabulary Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy Sales of Children’s Abilities - Verbal</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCarthy - Laterality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford-Binet IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>WISC III</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Edinburgh Handedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral Narratives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School Achievement Tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Behavioral Testing

<table>
<thead>
<tr>
<th>Test Type</th>
<th>8 Yrs</th>
<th>9 Yrs</th>
<th>10 Yrs</th>
<th>11 Yrs</th>
<th>12 Yrs</th>
<th>13 Yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERPs</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>WRAT Reading</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Woodcock-Johnson Reading Mastery Test - Word Attack</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Woodcock-Johnson Reading Mastery Test - Word Identification</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Phonological - Pronounceable Non-Words</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Spelling - identify correctly spelled word</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Orthographic - possible words given spelling rules</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>School Grades</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>School Achievement tests</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Longitudinal Research

😊 Pitfalls...

- New tests become “gold standards” in the field that are different from those you started with (new Bayley, new Stanford-Binet, new WISC)

- Children become more interactive with age which allows different types of assessments to be given (“point to this before that...” “tell me”)
  - Different tests can be used to reflect new abilities

- Children (or parents) can become familiar with the assessments

- Hazards in changing instruments!
SBIV versus WISC III

Ages of Children

Low
Medium
High

Intelligence Scores
Getting Started

Regardless of the design...

- You have to start somewhere
- The literature, colleagues, and inspection of assessment instruments can be helpful/critical (the Gold Standards)
- Need to choose behavioral tests that fit the model or constructs under study
- Need to choose tests with known validity and reliability, don’t “develop a scale”
Behavioral Testing In Infancy

Gold Standard Assessments
- Tests that are commonly used and well validated, reliable, and useful
- Bayley Scales of Infant Development (Bayley, 1993; new version)
  - Mental Development Index
  - Motor Development Index
  - Infant Behavior Record
Bayley Scales of Infant Development

- Range of ages is from 1 to 42 months

- Administered in age-based item sets of about 20 - 25 items

- Must get credit for a certain number of items to stay within a set or to move up or down a set (baseline, ceiling)
Basal and Ceiling Determination

- **Basal**: Assumes that child can perform all items below an item set.
- **Ceiling**: Assumes that the child cannot perform any items above an item set.
- Establishes where to start and where to end.
Testing An Infant
Bayley Scale

Mental Development Index (MDI): includes cognitive, language, social, and overall index score of cognitive skills

Can derive items related to language only for research purpose.

Can relate language items to performance on other language assessments - MacArthur Communicative Development Inventory (vocabulary, gestures, games and imitation), which is parental report

A multi-method approach
Preschool Assessments

General Assessments:

Stanford-Binet IV (or V) Intelligence Test: four areas, with a variety of subtests
- Verbal Reasoning (vocabulary, verbal comprehension, absurdities, verbal relations)
- Abstract/Visual Reasoning (pattern analysis, copying, matrices, paper folding and cutting)
- Quantitative (quantitative, number series, equation building)
- Short-Term Memory (bead memory, memory for sentences, memory for digits, memory for objects)
Stanford-Binet IV

- For preschool children
  - Only a subset of the subtests are given based on age and abilities
  - Factor structure does not support a four factor model, only two factor (verbal & nonverbal) or three factor (verbal, nonverbal, memory)
  - Yields subtest and a composite score
Stanford-Binet IV

Verbal Reasoning

- Vocabulary - picture and oral vocabulary
- Comprehension - follow directions, figure out what to do if, what are ____ for, etc.
- Absurdities - why is this odd? (tree upside down)
Preschool Behavioral Testing

Differential Ability Scales (Elliott, 1990)

- ages 2 through 17
- verbal, nonverbal (fluid reasoning), nonverbal/spatial, achievement, and special diagnostic tests
- Different subtests for lower preschool, upper preschool, and school-age levels
Example subscales

- Block Building - arrange blocks to match a model; remember exact block orientations
- Verbal Comprehension - follow directions, “give me….” (the car, the horse, all the yellow chips)
- Picture Similarities - match a card to one of an array of pictures
- Naming Vocabulary - picture vocabulary and brief naming vocabulary
Preschool Behavioral Testing

**NEPSY (Kirkman, Kirk & Kemp, 1997)**
- neuropsychological test
- 3-12 year
  - Assesses Attention and Executive Functions, Language and Communication, Sensorimotor, Visuospatial, Learning and Memory
  - used with typically developing children and children diagnosed with learning disabilities, attention-deficit/hyperactivity disorder, autism, and speech and language impairment
NEPSY

Example subtests

- Phonological Processing: requires the child to identify one of three pictures after the examiner says a segmented word
- Verbal Fluency: assesses the child’s ability to generate words in each of two categories: “animals” and “food or drink”
- Sentence Repetition: child repeats increasing longer sentences and phrases
- Early Number Concepts: counting, number concepts, and simple arithmetic
Important Considerations

Tests are constructed for specific purposes: read the directions!

💜 CANTAB - NP Executive Function Tasks (Tower of London, Memory Span).

- Limitations -
  - it is all computerized, that can work well with some populations, but not others
  - some tasks are quite long
  - the stimulus presentation displays were made for adults, children may find them boring
  - psychometrics are unknown for young children
Important Considerations

- **Child Behavior Checklist (Achenbach)** -
  - Limitations of use with typically developing children
    - a problem-based scale, so it may not be sensitive to variations in attention or activity level in children without behavior problems
    - Achenbach created his scoring system and program such that scores below 50 (the expected sample mean), are reported as 50, so 1/2 of the distribution is truncated (undifferentiated)
Assessments for School-Aged Children

- There is a greater variety of assessments for older children
- Many options for different types of assessments of the same constructs
  - For example, the “language” construct (vocabulary, passage comprehension, writing, grammar)
- Can assess children and adults (parents)
Wide Range Achievement Test
WRAT 3 (Wilkinson, 1993)

⚠️ Ages 5.0-74.11

⚠️ Purpose: To assess the status of skill development in areas of reading, spelling, and arithmetic achievement
WRAT 3 cont..

3 Subtests on Blue and Tan forms

• Reading - recognizing and naming letters and pronouncing words out of context; the Alphabetic Principle
• Spelling - Writing name; writing letters to dictation.
• Arithmetic - counting, reading number symbols, solving oral problems, performing written computations
Woodcock - Johnson III
(Woodcock & McGrew, 2001)

Ages 2.0 through Adult

Purpose: To measure intellectual abilities and academic achievement

Computerized scoring and profiling -

This scoring has good and bad aspects
Woodcock- Johnson III Tests of Achievement (WJ-III ACH)

22 tests measuring 5 curricular areas:
Peabody Picture Vocabulary Test  
(Dunn & Dunn 1997)

Ages 2.5 - 90+

Purpose- Measures an examinee’s receptive hearing/ vocabulary acquisition

Administered using picture plates of drawings to depict definitions
The task of the test taker is to select the best 1 of 4 pictures that represents the meaning of the stimulus word presented orally by the examiner.

- Word difficulty ranges from bus to nautical; concrete to abstract, nouns to verbs/adverbs.
Expressive Vocabulary Test
(Williams 1997)

Ages 2.5 - 90+

Purpose: Measures oral/spoken vocabulary knowledge with 2 types of items: Labeling and Synonym.
The first 38 items are labeling items. The examiner points to a picture or part of the body and asks a question. The examinee is asked for a one word answer that is a noun, verb, or adjective.

The remaining 152 items are synonym items. The examiner presents a picture and a stimulus word. The examinee responds with a one word answer that is a noun, verb, adjective, or adverb.
PPVT and EVT: The Perfect Couple

The PPVT and EVT are co-normed, which allows direct comparison of expressive and receptive vocabulary.

In both tests, examinees are administered only a subset of items that most closely approximate their ability level through the use of age appropriate start items and basal/ceiling rules.
Important Considerations

- It is important to know what the test is able to do, how it has been used and with whom, and what the limitations are.

- Test selection can be a fatal flaw in research and especially in applications for research funding.

- Outdated, dubious validity/reliability, insensitive,
Screening Tests

Sometimes just a simple assessment will work

- Characterize a population (developmentally delayed, normally developing)

- Test a specific skill area (receptive vocabulary, phonological processing)

- Screening for presence/absence of a skill
Behavior Screening Test

 понравился

There are not many reading tests for 3 and 4 year olds

– Get Ready To Read - Whitehurst & Lonigan (2001). Taps:
  • Phonological skills
  • Letter/sound/word recognition
  • Literacy (book knowledge, writing conventions)
  • http://www.getreadytoread.org/online_games.php
Get Ready To Read

Item Sample These pictures are: boy, fish, apple, car. Which one is car? Find car. If child answers incorrectly: That was a good try, but this is car. Let’s try again. Which one is car?

Item 10 Find the letter that makes a buh sound.
Get Ready To Read

Item 12
Some children wrote their name. Find the one that is written the best.

Item 16
This is ball, and these pictures are: zebra, shoe, wall, leaf. Does ball sound like zebra, shoe, wall, or leaf? Find the one that rhymes with ball.
Get Ready To Read versus Reading-Related Skills

• Print knowledge (understanding of books, printed letters and words)

• Emergent writing (text knowledge)

• Phonological sensitivity (letter sounds, rhyming, segmenting words).
For 3 year olds

- GRTR measures were useful in identifying low scoring children (sensitivity)
  - Vocabulary Measures (PPVT and EVT)
  - Phonological Awareness (phonological processing and rhyme detection)
  - Not Environmental Print

- Average GRTR score = 7.9 (out of 20)
For 4 Year Olds

For 4 year olds - these measures were useful in identifying low scoring children
- Language Measures (PPVT and EVT)
- Phonological Awareness (rhyme detection, phonological awareness)
- Not Environmental print

Average GRTR score = 11.87 (out of 20)
Proof of the Pudding

For 4-year-olds

- GRTR assessment in the fall
  - correlated with fall WRAT letter and word scores ($r = .52^{**}, .01$, respectively)
  - Correlated with spring WRAT letter and word scores ($r = .68^{**}, .25^{*}$, respectively)
Coordinating With ERP Testing

- Our assessments must fit with the needs of the ERP testing group, and vice versa.

- Testing time tolerated by child for ERP testing and for behavior testing is limited and is related to age.

- Sometimes the behavioral testing can “warm up” a child to the ERP testing situation.
Coordinating With ERP Testing

- **Warming up - nonverbal tasks** (pointing to pictures rather than naming them, block building), and simple puzzle forms

- **Head measurements** can be taken while children are engaged in some of these tasks if this is research, not a clinical assessment, and is needed to fit into a testing window

- **Activities during ERP testing** that don’t require instructions or verbal responses can be used
Behavior Testing

➔ Adds qualitatively to the information gained from ERP testing
  - answers the “yeah, .... but what is really happening?” question

➔ Useful for characterizing the participants and their skills

➔ Represents a multi-method approach to important research questions
The Next Frontier?

Capturing all the variance not accounted for by ERP testing, which can be sizable.
Questions?

dlmolfese@mac.com