Standardized Tests-a participative workshop

Alexis Davis, BScOT
Tanja Mayson MSc, BScPT
Sue Stewart MRSc, BScPT
Audio available - BSID-II, PDMS-2, Sensory Profile, MABC-2
No audio - AIM.BOT2, DTVP2, GMFM
Objectives

- To outline the reasons for using standardized tests
- To review test terminology
- To introduce eight different standardized tests
  - To describe when they might be used
  - To describe their psychometric properties
- To enable participants to discuss four tests more fully with a clinical expert.
Purpose?

- Why do we use standardized tests?
  - Can all tests predict later function?
  - Can all tests measure change?
  - Need to determine the reason for using the test and then decide which test is best.
Before Selecting a test....

You need to know
- Type of test
- Purpose of the test
- Population it serves
- Time to Complete
- Equipment Needed
Types of Tests

- **Norm-referenced tests** = Compare child with so called “norm” or average of a group of similar children
- the purpose is to discriminate

Examples??
Types of Tests

- **Criterion-referenced tests** = Compare performance of a child in terms of a set of criterion
- the purpose is to evaluate

Examples???
Purpose of a test

- Why are you going to use the test?
  - Discrimination
  - Prediction
  - Evaluation
  - Planning/Description
Terminology

- Chronological age
- Corrected age
- Basal score
- Ceiling score
Test Scores and Characteristics

- Raw score
- Standard score
- Percentile
- Age equivalent
- Developmental quotient
Reliability

- The idea that the test will get the same values if repeated.
- Actually a very complicated concept and there are many measures to ensure good reliability.
Reliability

- To be reliable
  - Provide consistent values with small errors of measurement (Absolute Reliability)
  - Be capable of differentiating between clients with whom the measure is being used (Relative Reliability).
Validity

- The idea that the test is actually testing what it says it is.
- A measure is valid to the extent that it measures what it is intended to measure. Validity implies that a measurement is relatively free from error i.e. a valid test is also reliable.
Characteristics of diagnostic and screening tools

- **Sensitivity (80%)**
  - Test’s ability to get a real positive

- **Specificity (90%)**
  - Test’s ability to get a real negative

- **Positive predictive value**
  - Estimate that positive result means positive

- **Negative predictive value**
  - Estimate that negative result means negative
How to report findings

- Extremely important to think about what the best way to “message” the results of your test.

- Case example
What kinds of things do we need to be cognizant of when using standardized tests?
We need to appropriately set the context for the testing situation

For the child
- Making them feel comfortable
- Letting them know what is going on

For the family
- Discussing expectations and reasons for the visit
The Context

- Standardized tests only give us part of the picture
The Context

- Standardized tests only give us part of the picture
  - The art and the science
  - Clinicians need to use their experience and clinical reasoning to put all of the pieces together.
Difficulties

- How many of us modify the administration of the tests?
Difficulties

- What should you do if you really feel you need to modify a measure?
Tests
The Bayley Scales of Infant Development, 3rd Edition (BSID-III)
BSID-III

- Purpose
  - to identify children with developmental delay
BSID-III

- **Type**
  - Norm-referenced

- **Purpose**
  - Discrimination, planning

- **Population**
  - Children at risk for developmental delay
BSID-III

- Age
  - 0 - 42 months

- Time to Complete
  - 50 - 90 minutes

- Equipment Needed
  - BSID-III kit, table, chair, floor space for motor items, stairs
BSID-III

Purpose of revision

- Update the normative data
- Develop 5 distinct scales
- Strengthen the psychometric quality
- Improve the clinical utility
BSID-III

- Purpose of revision
  - Simplify administration
  - Update item administration
  - Update stimulus materials
  - Maintain basic qualities of the Bayley Scales
BSID-III

- Cognitive subscale
- Language scale
  - Receptive language subscale
  - Expressive language subscale
- Motor scale
  - Fine motor subscale
  - Gross motor subscale
BSID-III

- Who can administer?
BSID-III

- Standardized Sample
BSID-III

- Reliability
- Validity
- Further considerations
Cases

- What types of children do you think this test would be most appropriate for?
- Should it be used for children with cerebral palsy? Down’s syndrome?
- What is the age span that you think clinically is most appropriate?
The Peabody Developmental Motor Scales, 2nd Ed. (PDMS-2)
Purpose

- To estimate a child’s motor performance in comparison to peers
- To determine discrepancies between gross and fine motor abilities
- To assist in goal development
- To evaluate progress
- To study motor development and ability in children as well as the effectiveness of motor interventions
PDMS-2

- Type
  - Norm and criterion-Referenced
- Purpose
  - Discrimination, evaluation, planning
- Population
  - Children suspected of having motor difficulties
PDMS-2

- Age
  - 0 – 71 months

- Time to Complete
  - 45 – 60 minutes for entire test; 20-30 minutes per subscale

- Equipment Needed
  - Kit, desk, two chairs, stopwatch, manual, and floor space for motor items
PDMS-2

- Improvements
  - Updated normative data
  - Normative data that is representative of the US population
  - Stratification by age of normative data
PDMS-2 (improvements)

- Revisions to administration and scoring format
- Elimination of items criticized in previous edition
- Revision and reformatting of activity cards into a motor activities program
- Evaluation of items using conventional and newer statistical analyses to better identify biased items
PDMS-2 (improvements)

- Computed reliability coefficients for subgroups of normative sample
- Completion of new validity studies
- Availability of more specific performance levels for the scoring criteria
- Addition of illustrations to assist with administration
PDMS-2

- Gross motor scale
  - Reflexes
  - Stationary
  - Locomotion
  - Object manipulation
- Fine motor scale
  - Grasping
  - Visual motor integration
PDMS-2

Who can administer?
PDMS-2

- Standardized Sample
PDMS-2

- Reliability
- Validity
- Further Considerations
Cases

- What types of children do you think this test would be most appropriate for?
- Should it be used for children with cerebral palsy? Down’s syndrome?
- What is the age span that you think clinically is most appropriate?
The Sensory Profile

Purpose

- To evaluate the contributions of sensory processing to a child’s daily functional performance.
- To determine the child’s tendencies to respond to stimuli.
- To understand which systems are likely contributing or providing challenges to the child’s performance.
The Sensory Profile

- **Type**
  - Norm-referenced

- **Purpose**
  - Discrimination

- **Population**
  - Children at risk of sensory processing difficulties
The Sensory Profile

- **Age**
  - 0 to 3 years (Infant/Toddler Sensory Profile), 3 to 10 years (Sensory Profile), 11 years and up (Adolescent/Adult Sensory Profile)

- **Time to Complete:**
  - 30 minutes for Sensory Profile (10 minutes for Short Sensory Profile); 15 minutes for Infant/Toddler Sensory Profile; 10-15 minutes for Adolescent/Adult Sensory Profile

- **Equipment Needed**
  - Questionnaire, manual, score form, pen/pencil
The Sensory Profile

1. The Sensory Profile (SP)
2. The Infant/toddler Sensory Profile (ITSP)
3. The Adolescent/Adult Sensory Profile (AASP)
The Sensory Profile

1. The Sensory Profile (SP)
   1. Sensory processing
      1. Targets child’s responses to basic sensory processing systems
   2. Modulation
      1. Reflects child’s regulation of neural messages
   3. Behavioral and Emotional Responses
      1. Reflects child’s behavioral outcomes of sensory processing
Who should administer?
The Sensory Profile

- Standardized Sample
  - SP
  - ITSP
  - AASP
The Sensory Profile

- Reliability
- Validity
- Further Considerations
Cases

- What types of children do you think this test would be most appropriate for?
- Should it be used for children with substance exposure? Cerebral palsy?
- What is the age span that you think clinically is most appropriate?
Movement Assessment Battery for Children (MABC-2)
MABC-2

- Purpose
  - Identification of movement difficulties
  - Clinical exploration and intervention planning
  - Program evaluation
  - Research
MABC-2

- **Type**
  - Norm-referenced

- **Purpose**
  - Discrimination, planning, evaluation

- **Population**
  - Children with motor difficulties
MABC-2

- **Age**
  - 3 – 16 years

- **Time to Complete**
  - 20 – 40 minutes for test

- **Equipment Needed**
  - MABC-2 kit, stopwatch, table, two chairs, floor space for motor items
MABC-2

- A revision of test content
  - New plastic materials
  - Improving certain task items
  - Clarifying instructions

- A revision of test structure
  - Extending ages to 3 to 16 years
  - Reduction to three age bands
MABC-2

- Who can administer?
MABC-2

- Standardized Sample
MABC-2

- Reliability
- Validity
- Further Considerations
Cases

- What types of children do you think this test would be most appropriate for?
- What is the age span that you think clinically is most appropriate?
Agenda

- No audio accompaniment – AIM.BOT2, DTVP2, GMFM
Alberta Infant Motor Scale (AIMS)

Purpose is to measure motor development in high risk infants
AIMS

- Type
  - Norm referenced

- Purpose
  - Discrimination
  - Evaluation

- Population
  - At risk infants
AIMS

- **Age**
  - 0-18 months

- **Time to complete**
  - 20-30 minutes

- **Equipment needed**
  - Book, form, pen and observation surface
Overview
- 58 item observational assessment
- observed or not observed
- scores are added up
- higher score means more mature
- percentiles and age equivalent
- cut offs?
AIMS

- Who can administer?
AIMS

- Standardized Sample
AIMS

- Reliability
- Validity
- Further considerations
Cases

- What types of children do you think this test would be most appropriate for?
The BOT-21 (2005) is a revision of the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) (1978).
BOT2

<table>
<thead>
<tr>
<th>Purpose</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Support diagnoses of motor impairments</td>
<td></td>
</tr>
<tr>
<td>Screen for motor impairment and determine the need for further assessment or intervention</td>
<td></td>
</tr>
<tr>
<td>Make placement decisions regarding physical education programs</td>
<td></td>
</tr>
<tr>
<td>Develop and evaluate motor training programs</td>
<td></td>
</tr>
<tr>
<td>Assist clinicians and researchers</td>
<td></td>
</tr>
</tbody>
</table>
BOT-2

- Type: Norm-referenced
- Purpose: Discrimination, planning, evaluation
- Population: Children at risk for motor impairment
BOT-2

- **Age:**
  - 4 – 21 years

- **Time to Complete:**
  - 40 – 60 minutes for complete form; 15 – 20 minutes for short form

- **Equipment Needed:**
  - BOT-2 kit, stopwatch, table, two chairs, floor space for motor items
BOT-2 changes

- Improve functional relevance
- Expand coverage
- Improve measurement among 4 and 5 year olds
- Extend norms through age 21
- Improve item presentation
- Improve quality of kit equipment
BOT-2

- Administration
- 4 ways to administer
BOT-2

- Standardized Sample
BOT-2

- Reliability
- Validity
- Further considerations
Cases

- What types of children do you think this test would be most appropriate for?
- Should it be used for children with learning difficulties?
- What is the age span that you think clinically is most appropriate?
The Developmental Test of Visual Perception, 2nd Edition (DTVP-2)
Purpose

- Document the presence and degree of visual perceptual or visual-motor difficulties in individual children
- Identify candidates for referral to intervention programs
- Verify the effectiveness of intervention programs
- Serve as a research tool
DVTP-2

- **Type**
  - Norm-referenced

- **Purpose**
  - Discrimination, planning, evaluation

- **Population**
  - Children at-risk for visual perceptual or visual-motor difficulties
DVTP-2

- **Age**
  - 4 – 10 years

- **Time to Complete**
  - 30 – 60 minutes

- **Equipment Needed**
  - DTVP-2 kit, table, chair
The DTVP-21 (1993) is a revised version of the original DTVP2 (1966).

- An increase in reliability of subtests to acceptable levels
- Evidence for content, criterion-related, and construct validity
- Completion of factorial analysis
- Demonstration of an absence of racial, gender, and handedness bias
DVTP-2

- Inclusion of a normative sample whose demographic characteristics are similar to the US population
- Development of two new composite scores (motor-reduced visual perception and visual-motor integration) to assist with diagnosis
- Expansion of age tested to include 10 year olds
DVTP-2

Subtests
- Eye-hand coordination
- Position in space
- Copying
- Figure-ground
- Spatial relations
- Visual closure
- Visual-motor speed
- Form constancy
DVTP-2

- Administration
DVTP-2

- Standardization Sample
DVTP-2

- Reliability
- Validity
- Further considerations
Cases

- What types of children do you think this test would be most appropriate for?
- What is the age span that you think clinically is most appropriate?
The Gross Motor Function Measure (GMFM)
GMFM-88 and GMFM-66

- Purpose

  - to evaluate motor skills in children with cerebral palsy (CP).
GMFM

- **Type**
  - Criterion-referenced

- **Purpose**
  - Evaluation

- **Population**
  - Children with cerebral palsy (CP)
GMFM

- **Age**
  - 5 months to 16 years with motor skills at or below the level of a typical 5 year old.

- **Time to Complete**
  - 45 – 60 minutes for GMFM-88; less for GMFM-66

- **Equipment Needed**
  - Mat, bench, toys, and access to five stairs
GMFM

- Administration
GMFM

- Standardized Sample
GMFM

- GMFM 66 (advantages)
  - Ordering of items according to difficulty
  - Change to an interval scale which allows for better interpretability
  - Decrease in administration time
  - Use of a computer scoring system
GMFM

- GMFM 66 (disadvantages)
  - Less items in the lower dimensions which may make the GMFM-66 less descriptive for children functioning at lower levels
  - Need for computer and software to score
  - Need to learn to interpret item maps
GMFM

- Reliability
- Validity
- Further Considerations
Cases

- What types of children do you think this test would be most appropriate for?
- Should it be used for children with cerebral palsy? Down’s syndrome?
- What is the age span that you think clinically is most appropriate?
Case Studies
References
