

# **E-411 PRMA**

## **LECTURE 16 - INTELLIGENCE TESTS & EDUCATION**

Christopher David Desjardins

## LAST TIME

Different models of intelligence

Spearman's  $g$ ; Cattell & Horn's broad domains; Carroll's hierarchical view; CHC; processing view of intelligence

Does it really matter how we define intelligence?

Is it more academic than practical?

# TEST SELECTION CRITERIA

- Theoretical model
- Ease of administration and scoring
- Ease on interpretation for a purpose
- Appropriateness of the norms
- Reliability and validity indices
- Test utility (i.e. costs vs. benefits)

# STANFORD-BINET IQ TEST

- Conceived to screen for developmental disabilities
- Provided organized administration/scoring instructions
- Originally intelligence calculated as the ratio of mental age to true age (IQ)
- Deviation IQ, mean of 100 and standard deviation of 16 (now 15), comparison of an individual with other of the same age in the standardization sample
- Used **age scale**, grouping by items by age; **point scale**, grouping of items by category that could be scored correct or incorrect
- Scores can be obtained as a **test composite** (e.g. IQ)

## FIFTH EDITION

- For 2 to 85+ year olds
- Measures fluid reasoning, knowledge, quantitative reasoning, visual-spatial processing, and working memory (table 10-2).
- 10 subtests, each about 5 minutes ( $M = 10$ ;  $SD = 3$ )
- Scores can be used to create various composites: factor index; nonverbal and verbal IQ; abbreviated battery IQ; and FSIQ

# USES OF SB5

- Diagnose a range of development disabilities and exceptionalities
- Clinical and neuropsychological assessment
- Early childhood assessment
- Psychoeducational evaluations for special education placements
- Adult social security and workers' compensation evaluations
- Providing information for interventions such as IFSPs, IEPs, career assessment, industrial selection, and adult neuropsychological treatment
- Forensic contexts
- Research on abilities and aptitudes

# VALIDITY EVIDENCE FOR SB5

Normative sample: 4,800 individuals between 2 and 85+ years (to match the 2000 U.S. Census)

Bias reviews were conducted on all items for the following variables: gender, ethnicity, culture, religion, region, and socioeconomic status.

Co-normed with the Bender Visual-Motor Gestalt Test, Second Edition and the Test Observation Form

For the FSIQ, NVIQ, and VIQ, reliabilities range from .95 to .98.

Reliabilities for the Factor Indexes range from .90 to .92.

For the 10 subtests, reliabilities range from .84 to .89.

Concurrent and criterion validity data were obtained using the SB-IV, SB-LM, WJ III, UNIT, Bender-Gestalt II, WPPSI-R, WAIS-III, WIAT®-II, and WISC-III.

# WISC-IV

Provides IQ scores and critical clinical insights into a child's cognitive functioning.

Measures verbal comprehension; perceptual reasoning; working memory; and processing speed (table 10-5)

Norming: consisted of 2,200 children between the ages of 6 and 16:11 years. The sample was stratified on age, sex, parent education level, region, and race/ethnicity.

What kind of validity evidence is provided?

- Psychometric report
- Clinical validity report

# MEASUREMENT CHECK

Will the reliability and validity statistics reported in a technical manual be applicable to you as a test administrator?

If they report coefficient alpha of 0.95, what will your coefficient alpha be?

If the correlation between the IQ score obtained from the Stanford-Binet and an GRE is reported as .75, will you have that same correlation for your group of students?

# COMPARISON AND OTHER TESTS

- Both purport to measure intelligence
- Highly correlated, differ by amount of  $g$
- Both work within the CHC model, though Weschler favors  $g$
- Both represent gold standard
- Kaufman test focus on processing not structure

# GROUP TESTS

- USA army developed tests for recruits in WWI
- Alpha, those who could read, Beta, those who couldn't
- Assigned duty and service based on performance
- Tests used in post-war because they were much cheaper
- Later, Army General Classification Test and Armed Service Vocational Aptitude Battery
- Also used in the schools in the USA for placement (not as much now)

**WHAT IS THE PURPOSE OF SCHOOL AND  
EDUCATION?**

**HOW DO WE USE TESTS IN EDUCATION?**

---

*In the United States, historically, the purpose of education has evolved according to the needs of society.*

*Education's primary purpose has ranged from instructing youth in religious doctrine, to preparing them to live in a democracy, to assimilating immigrants into mainstream society, to preparing workers for the industrialized 20th century workplace.*

---

---

*I think that my view, and most people's view, is that the purpose of education is to support children in developing the skills, the knowledge, and the dispositions that will allow them to be responsible, contributing members of their community –their democratically-informed community. Meaning, to be a good friend, to be a good mate, to be able to work, and to contribute to the well-being of the community.*

# ELEMENTARY AND SECONDARY EDUCATION ACT AKA NO CHILD LEFT BEHIND

- Educational opportunities for all students
- One measurement of NCLB, adequate yearly progress is a measurement used to determine how every public school and school district is performing according to results on a standardized exam.
  - Yearly progress towards 100% proficiency (student is working successfully on grade level material)
  - Proficiency measured by standardized assessment
  - Other criteria (e.g. attendance)

# MN AYP

- For Minnesota, in 2010, "Adequate Yearly Progress" (AYP) is measured by splitting students into 27 subgroups, based on the subject area and students' ethnicity. (Hispanic students' performance on math is one example). If students in just one of those subgroups don't meet proficiency benchmarks, the entire school is considered "Not Making AYP."
- [Minnesota Report Card](#)

# RESPONSE TO INTERVENTIONS

- **Specific learning disability**, 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, think, speak, read, write, spell, or to do mathematical calculations
- SLD diagnosis based on a **response to intervention**

# RTI

- Data driven
- Multilevel preventive approach
  - Regular instruction
  - Small-group instruction
  - Individual instruction
- Instruction, evaluation, intervention, reevaluation  
(repeat)
- Can be developed school-wide or tailored individually
- Multiple sources should be used with diagnosis (an **integrative assessment**)

# DYNAMIC ASSESSMENTS

- Based on a test-intervention-retest model
- Focuses on your learning potential or **zone of proximal developmental**
- Test administrator actively cheers you on and prepares you to be retested

What is the difference between an achievement and an aptitude test?

How do their uses differ?

Is it possible to write an item that measures achievement and not aptitude or vice versa?

# ACHIEVEMENT VS. APTITUDE TESTS

- Achievement test, measures degree to which you have acquired knowledge or skill, typically as a function of instruction or training.
  - Relatively defined learning experience
- Aptitude test, focus more on informal learning or life experiences. "Culmination of what you know".
  - Less defined

# ACHIEVEMENTS TEST AND USES

- May cover many academic areas and be divided into subtests (**batteries**)
- Norm, criterion-referenced, or both
- Batteries have used as practice tests, locate what test a student should take, co-normed with aptitude tests to allow comparison
- **Curriculum-based assessment** - measurements that use a student's performance on local curriculum as a basis for gathering information to make instructional decisions.
- **Must align with educational objectives and curriculum!**
- End of HR exam in psychology?

# TYPES OF QUESTIONS

- Fact-based items
- Conceptual items

# APTITUDE TESTS

- Prognosis or readiness tests
  - SAT, ACT, GRE, LSAT, MCAT, TIMSS, PISA
- Preschool level could include rating scale or checklists, like CBCL
  - Anxiousness, withdrawn, somatic complaints, social problems, thought problems, attention problems, rule-breaking behavior, and aggressive behavior
  - Connors for ADHD

# PISA

- Tests literacy in reading, mathematics, and science on a 1000 point scale
- <http://tinyurl.com/pbj53ad>

# **DIAGNOSTIC VS. EVALUATIVE PURPOSES**

**Evaluative information** - tests are used to make judgements

Pass or fail a course, program

Admission to a program

**Diagnostic information** - tests are used to identify

difficulties for remedial services

**diagnostic test**

# DIAGNOSTIC TESTS

- May consist of multiple subtests
- Designed to identify the missing knowledge/skill
- Typically, easier than evaluative tests
- Doesn't answer *why*
- Often focus on **reading** and **mathematics**

**WHAT WOULD AN ITEM ON A READING DIAGNOSTIC TEST  
LOOK LIKE?**

**WHAT WOULD AN ITEM ON A MATHEMATICS DIAGNOSTIC  
TEST LOOK LIKE?**

**HBJ** HARCOURT BRACE JOVANOVICH, INC. New York

Copyright © 1976 by Harcourt Brace Jovanovich, Inc. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Printed in U.S.A.

# WOODCOCK READING MASTERY

RED LEVEL FORM B  
TEST BOOKLET

## Stanford Diagnostic Reading Test

Bjorn Karlsen  
Richard Madden  
Eric F. Gardner

measures: reading readiness, achievement, and difficulties



**HBJ** HARCOURT BRACE JOVANOVICH, INC. New York

Copyright © 1976 by Harcourt Brace Jovanovich, Inc. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Printed in U.S.A.

RED LEVEL FORM B  
TEST BOOKLET

## Stanford Diagnostic Reading Test

Bjorn Karlsen  
Richard Madden

## STANFORD DIAGNOSTIC TESTS-REVISED Reading Test

Bjorn Karlsen  
Richard Madden  
Eric F. Gardner

norm: 3,300 USA nationally representative.

target: 4.5 to 80 year olds

subtests include letter identification, word identification,  
word attack, word comprehension, passage  
comprehension, phonological awareness, listening

**HBJ** HARCOURT BRACE JOVANOVICH, INC. New York

Copyright © 1976 by Harcourt Brace Jovanovich, Inc. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Printed in U.S.A.

RED LEVEL FORM B  
TEST BOOKLET

## Stanford Diagnostic Reading Test

Bjorn Karlsen  
Richard Madden

**HBJ** HARCOURT BRACE JOVANOVICH, INC. New York

Copyright © 1976 by Harcourt Brace Jovanovich, Inc. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Printed in U.S.A.

RED LEVEL  
TEST BOOKLET

## Stanford Diagnostic Reading Test

Bjorn Karlsen  
Richard Madden  
Eric F. Gardner



**HBJ** HARCOURT BRACE JOVANOVICH, INC. New York

Copyright © 1976 by Harcourt Brace Jovanovich, Inc. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Printed in U.S.A.

RED LEVEL  
TEST BOOKLET

## Stanford Diagnostic Reading Test

Bjorn Karlsen  
Richard Madden

**STANFORD DIAGNOSTIC READING  
TESTS**

**INSTEAD: GRADE**



**KEYMATH3**

# PSYCHOEDUCATIONAL BATTERIES

• used for normative comparison and to plan interventions

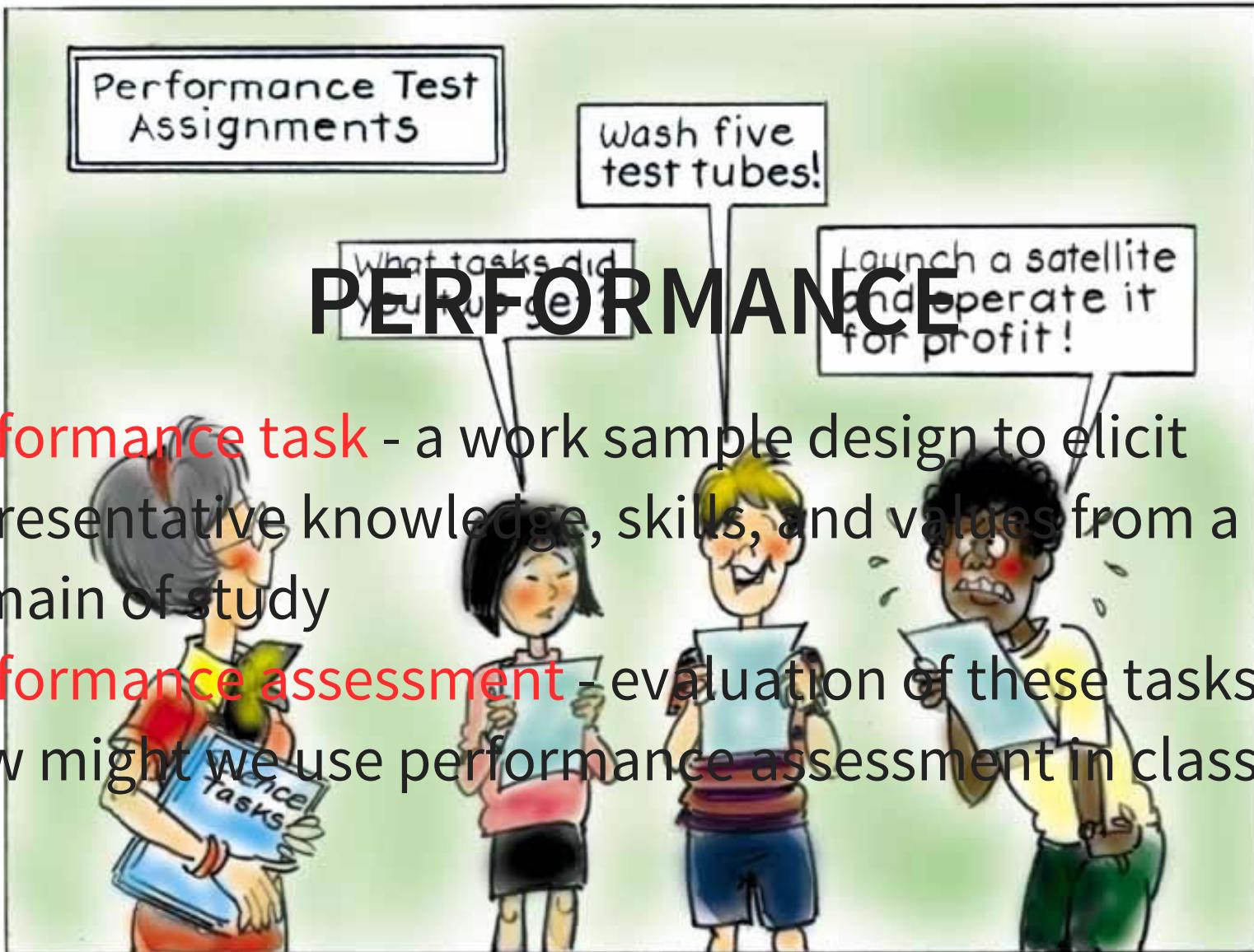
**HOW WOULD AN ITEM HERE DIFFER  
FROM THOSE FOR A DIAGNOSTIC  
TEST?**

# KAUFMAN ASSESSMENT BATTERY FOR CHILDREN

- Measures intelligence and achievement
- Kaufmans focus on information-processing aspect of intelligence
  - Simultaneous - all at once
  - Sequential - processing in a series
- Table 11 - 3
- Unclear factor structure
- Also can be used with the CHC model ... but how?

# WOODCOCK-JOHNSON IV

- Contains three batteries
  - Cognitive ability
  - Achievement
  - Oral language
- Administer together or individually
- [Assessment Service Bulletin](#)



- **Performance task** - a work sample design to elicit representative knowledge, skills, and values from a domain of study
- **Performance assessment** - evaluation of these tasks
- How might we use performance assessment in class? HR?

# PORTFOLIO

- What is a portfolio and what are some examples of a portfolio?
- A sample of your work
- How might we use a portfolio in class? HR?
- What are some ways you use a portfolio?
- Major issue, potential subjectivity in scoring

# AUTHENTIC ASSESSMENT

- A form of performance assessment is *authentic assessment*
- A task that evaluates your ability to transfer knowledge from the classroom to the real-world
- What have we done in class that is this type of an assessment?
- Major issue, could be affected by what you already know

# PEER APPRAISAL AND OTHER MEASUREMENTS IN EDUCATION

peers assign a score or ranking to you

"Which student would you rather work on a class project with?"

"Which student is the most popular"

these are often dynamic

Other inventories measure study habits, interests, and attitudes