

Item Response Modeling and its Applications

Margaret Wu

Educational Measurement Solutions / Victoria University

Report Card

- Amy

Subject	Mark
Mathematics	78
Reading	85
Science	75
Geography	80

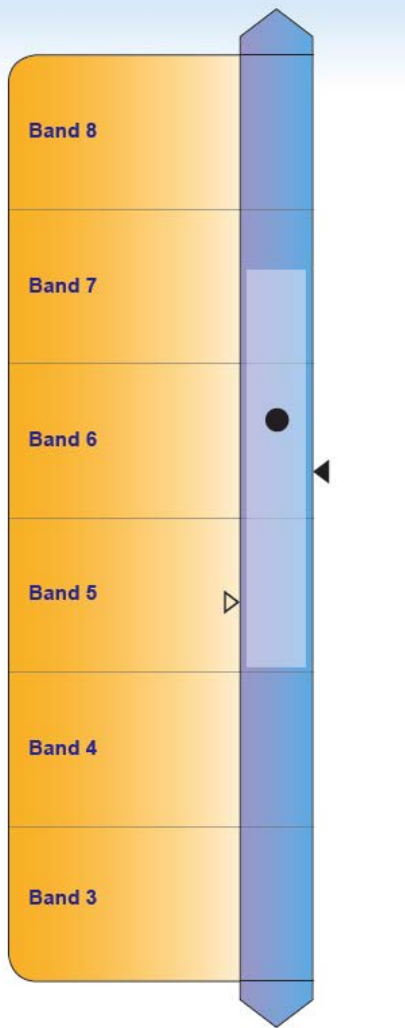
- Ben

Subject	Grade
Mathematics	A
Reading	B
Science	B+
Geography	C

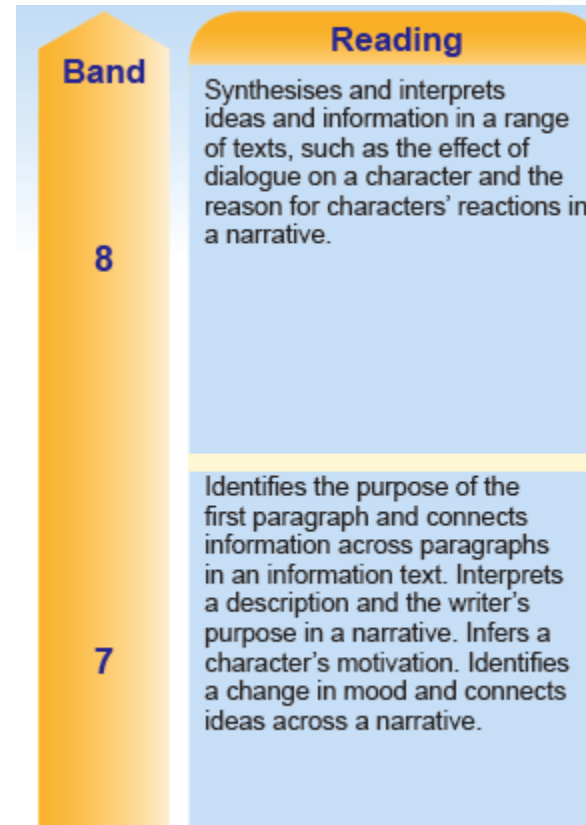
Issues

- Norm referencing – standing within a group?
- Criterion referencing - What can the students actually do?
- Comparability from one year to another
- Comparability from one student to another

Year 5



NAPLAN



KEY

● Individual student result

◀ National average

▷ School average

Range of achievement for the middle 60% of Year 5 students in Australia

Properties of an Ideal Measurement

- Scores we obtained are meaningful.



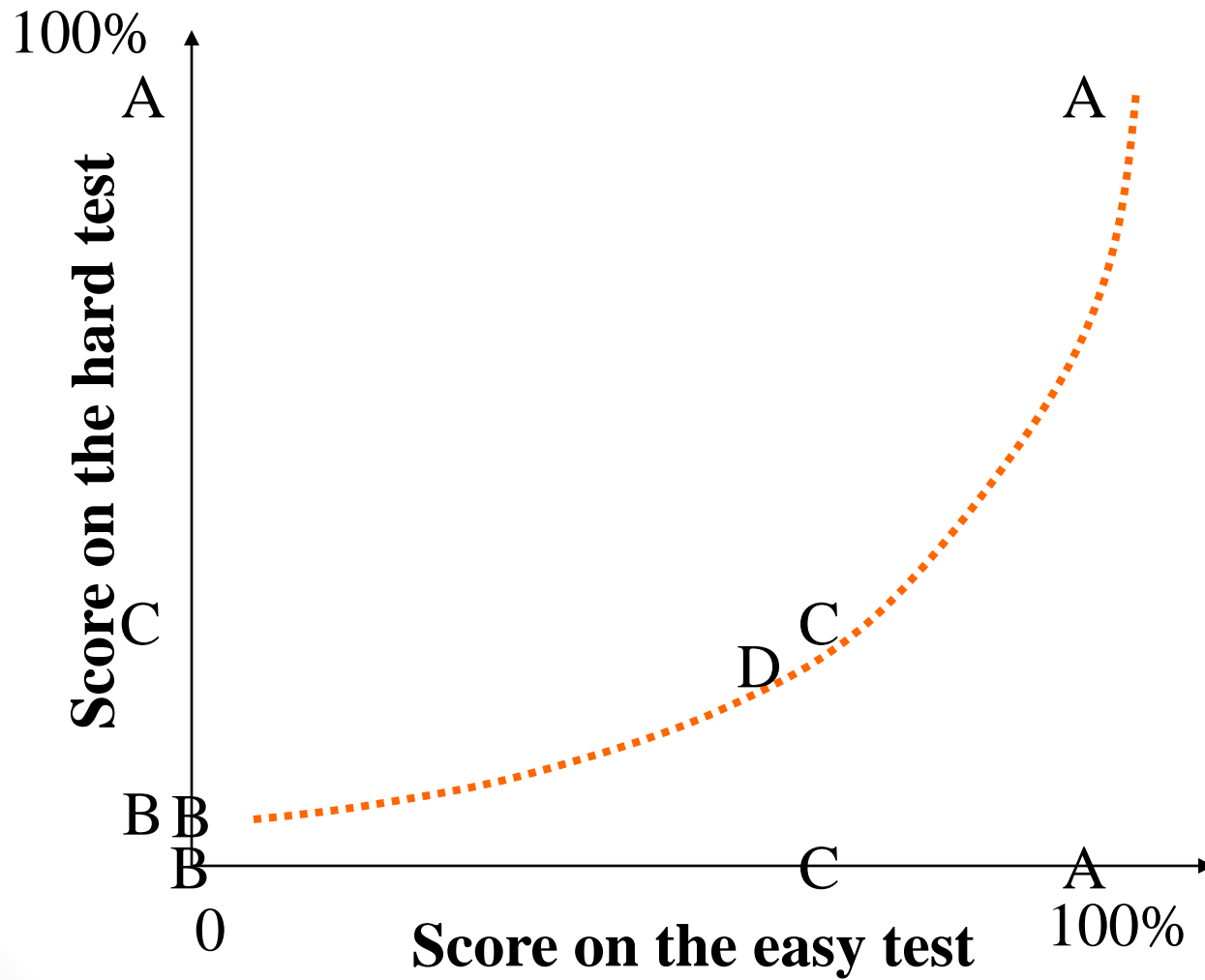
- What can each of these students do?
- Scores are independent of the sample of items used
 - If a different set of items are used, we will get the same results, in terms of the placement of the people on a scale.



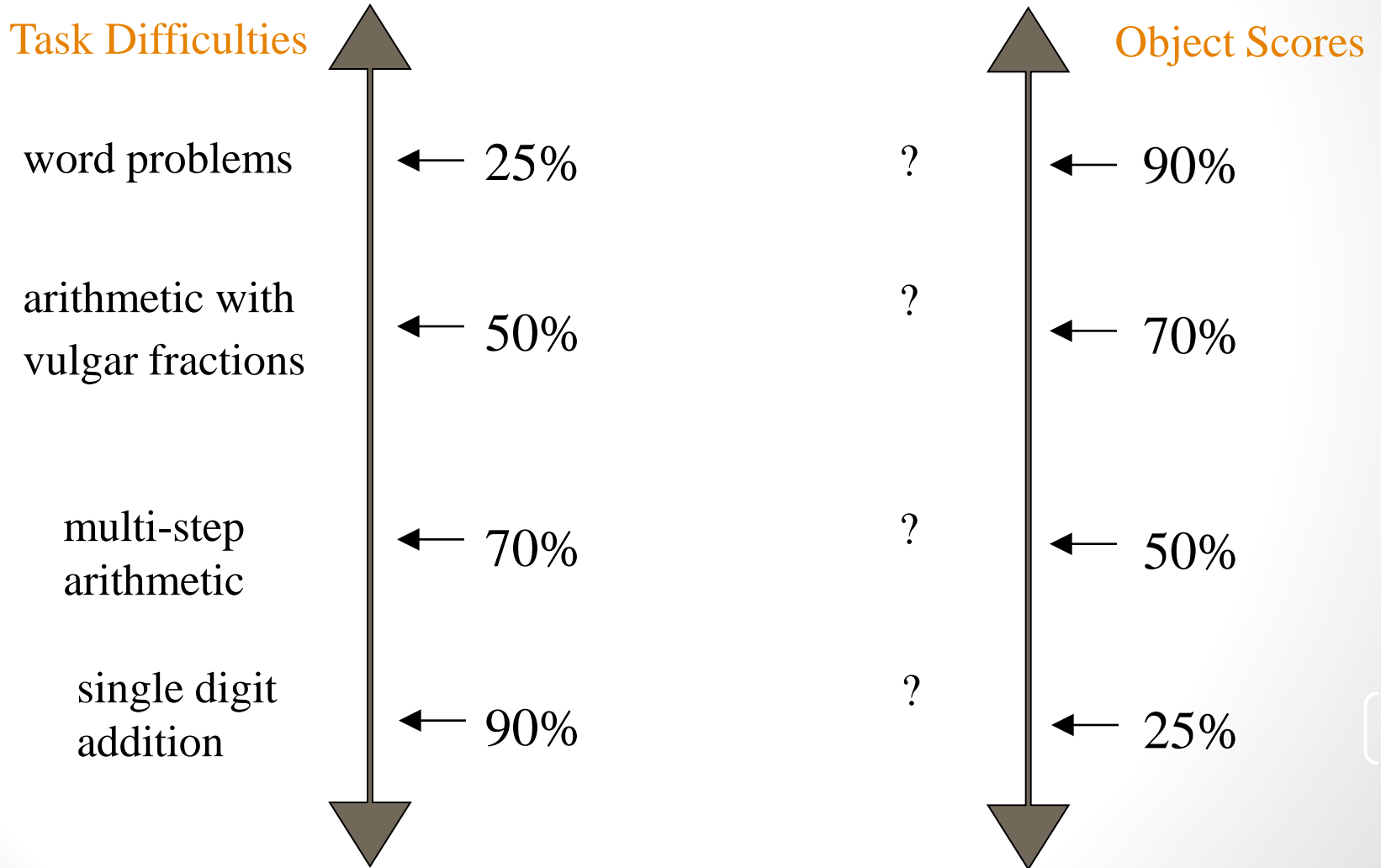
Using Raw Test Scores

- Can raw scores provide the properties of an ideal measurement?
- Distances between differences in scores are not easily interpretable.
- Difficult to link item scores to person scores.

Equating raw scores - 2



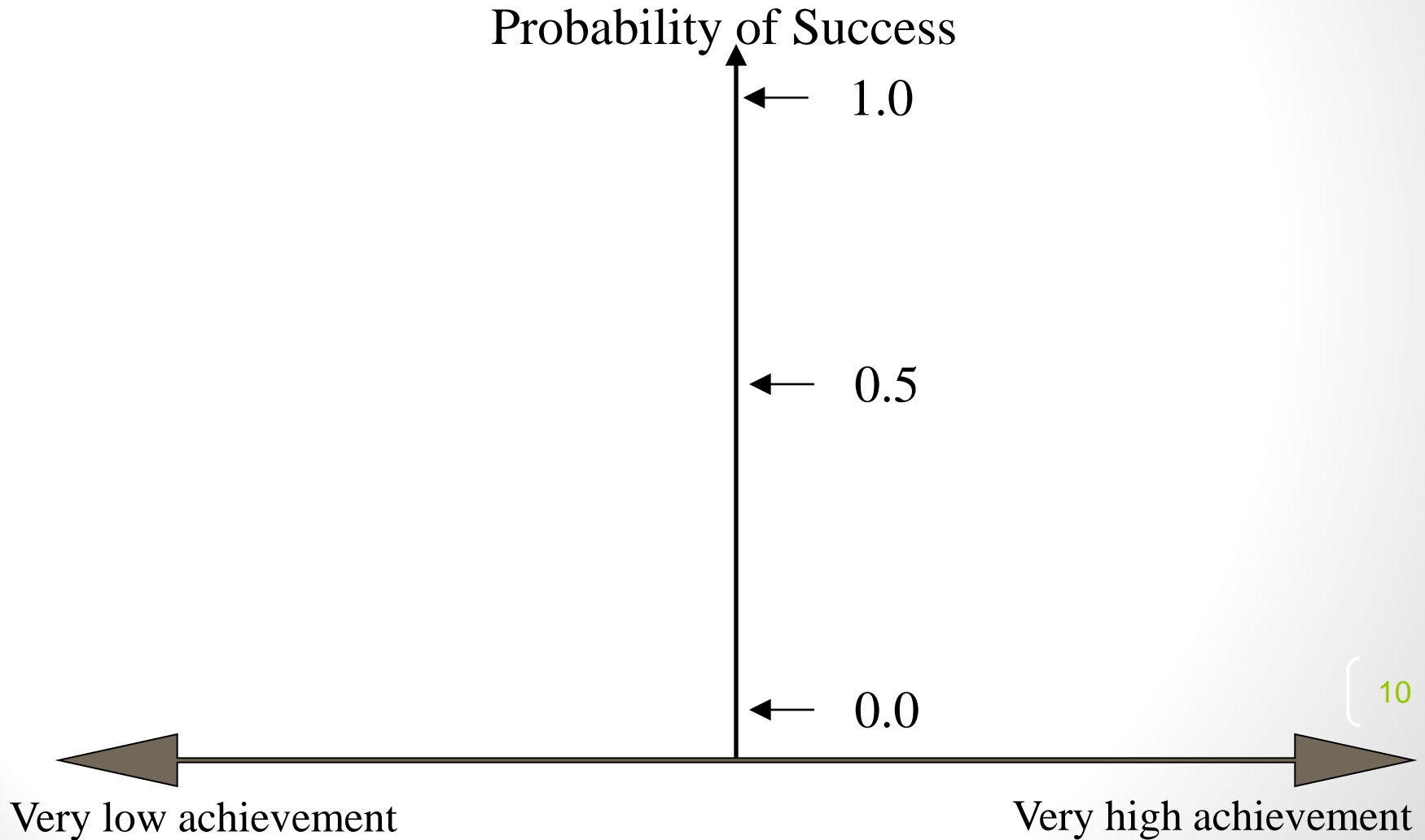
Link Raw Scores on Items and Persons



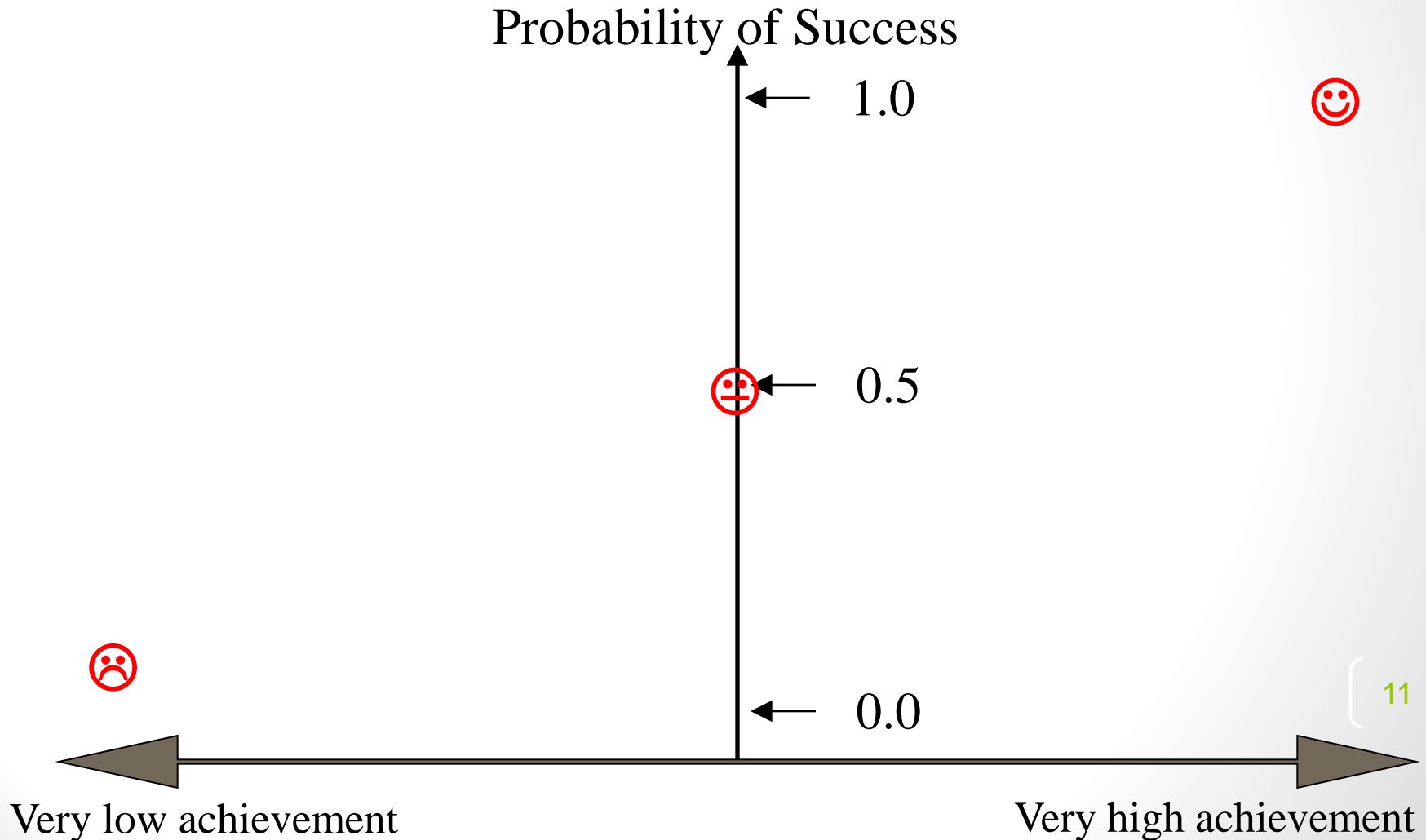
IRT

- IRT models give the *probability* of success of a person on an item.
- IRT models are not deterministic, but probabilistic.
- Given the “item difficulty” and “person ability”, one can compute the probability of success for each person on each item

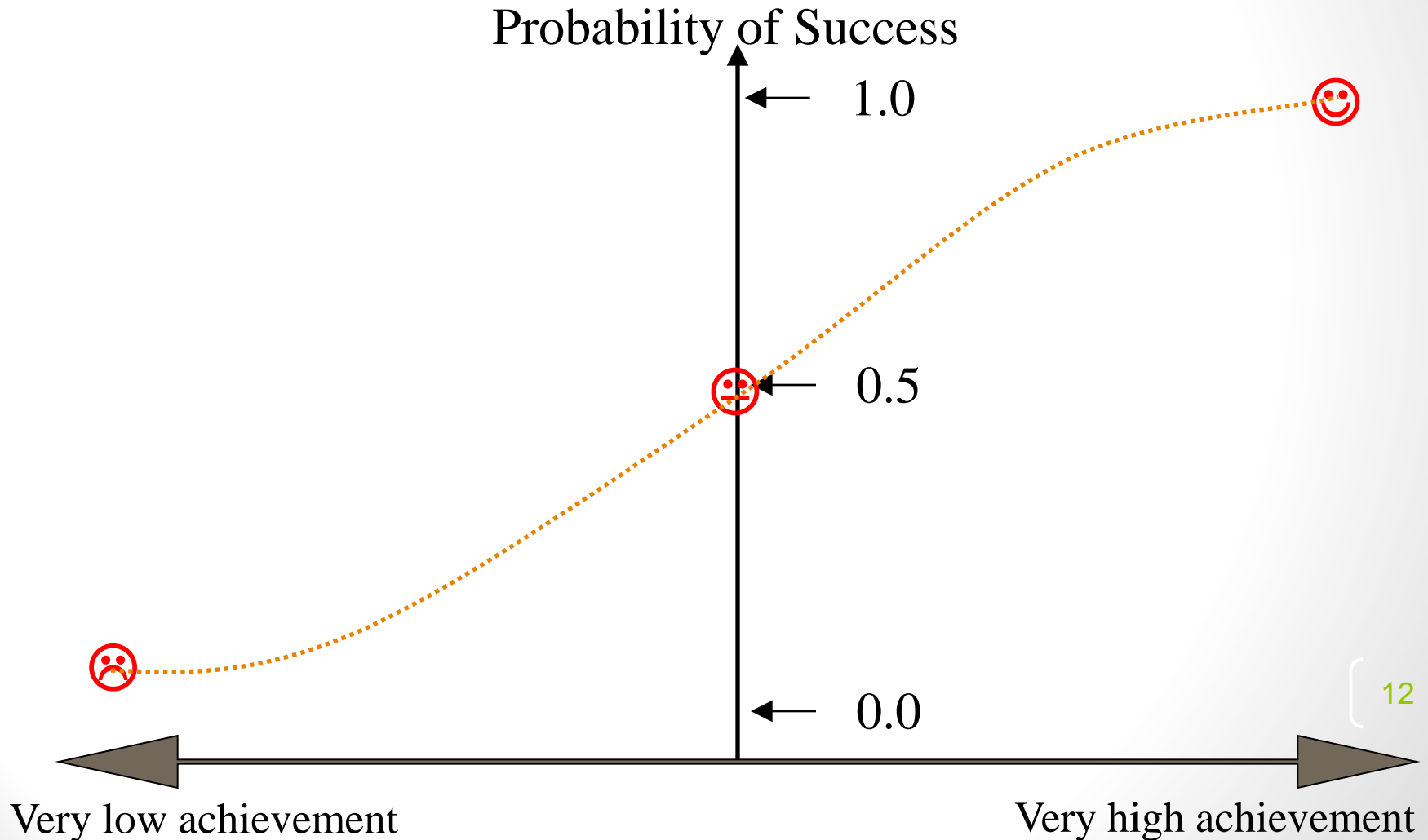
Building a Model



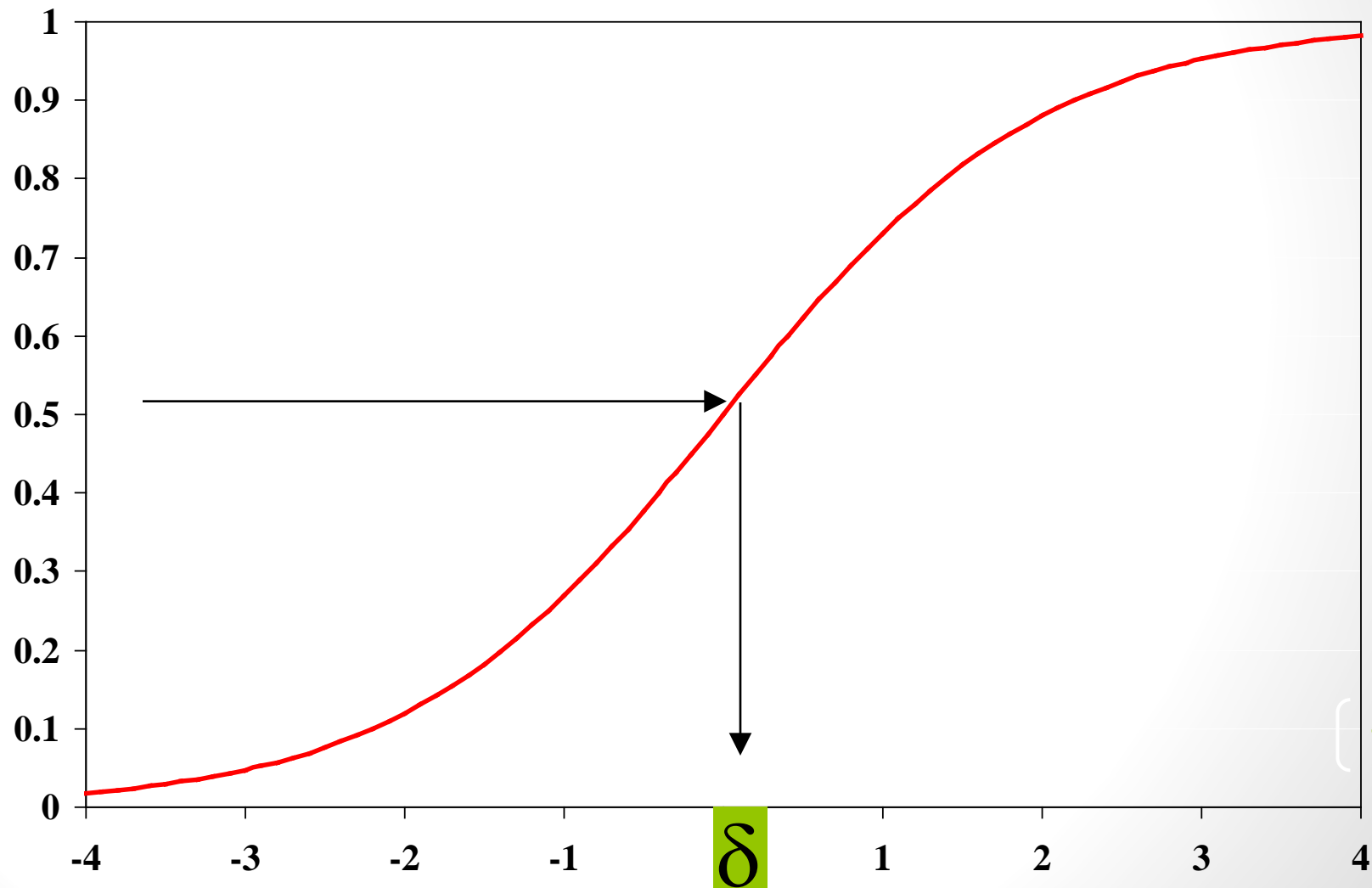
Imagine a middle difficulty task



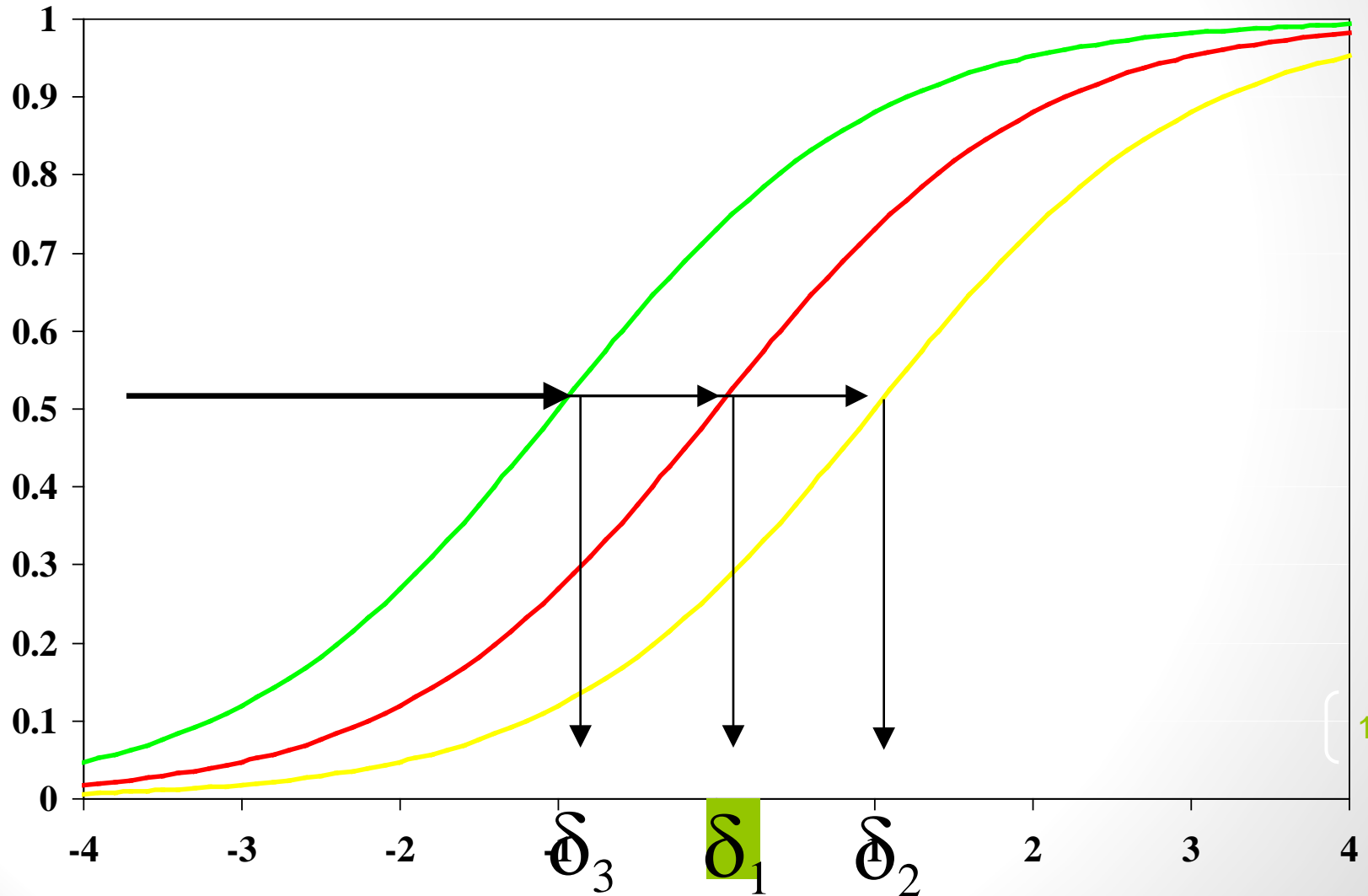
Item Characteristic Curve



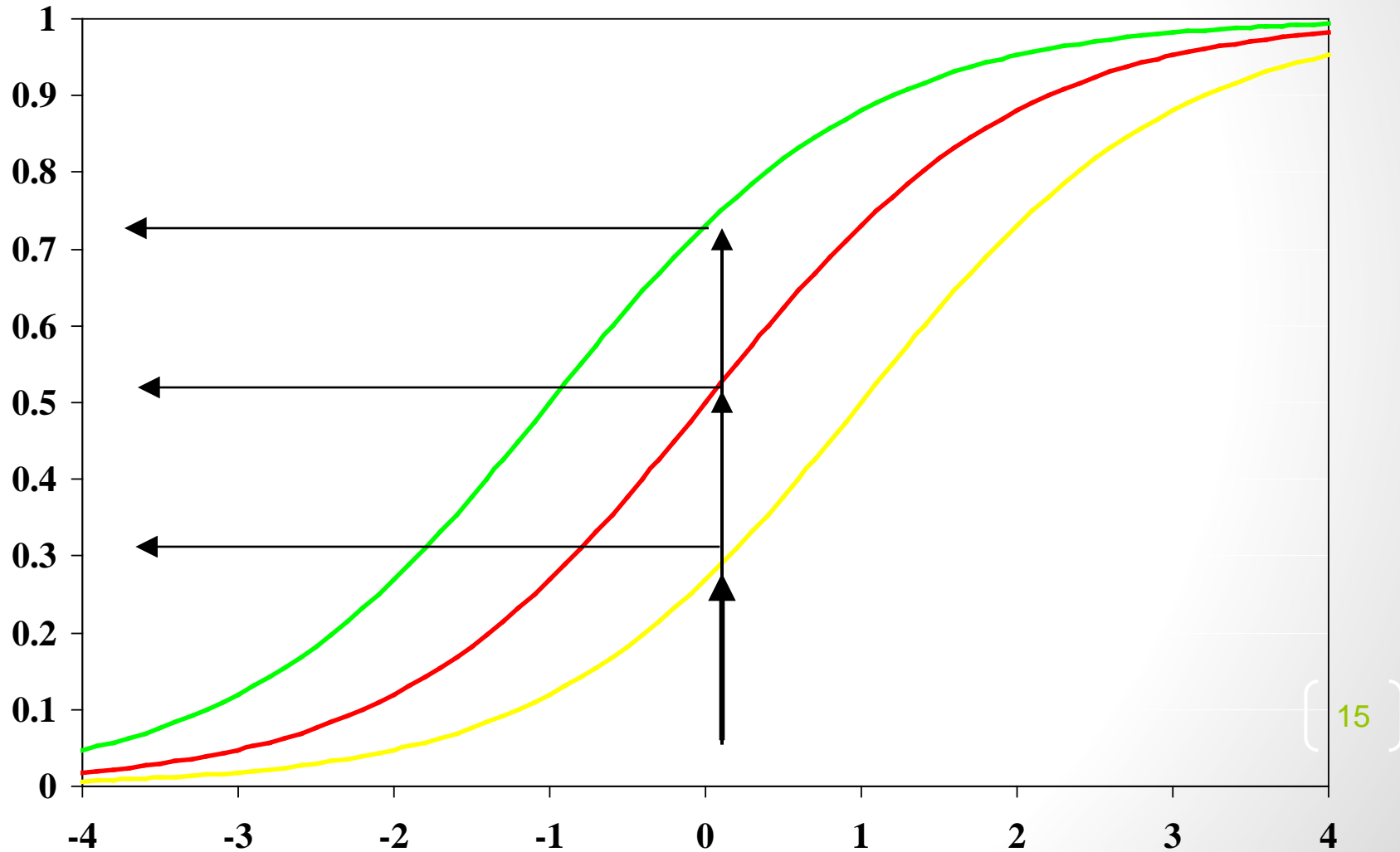
Item Difficulty -- 1



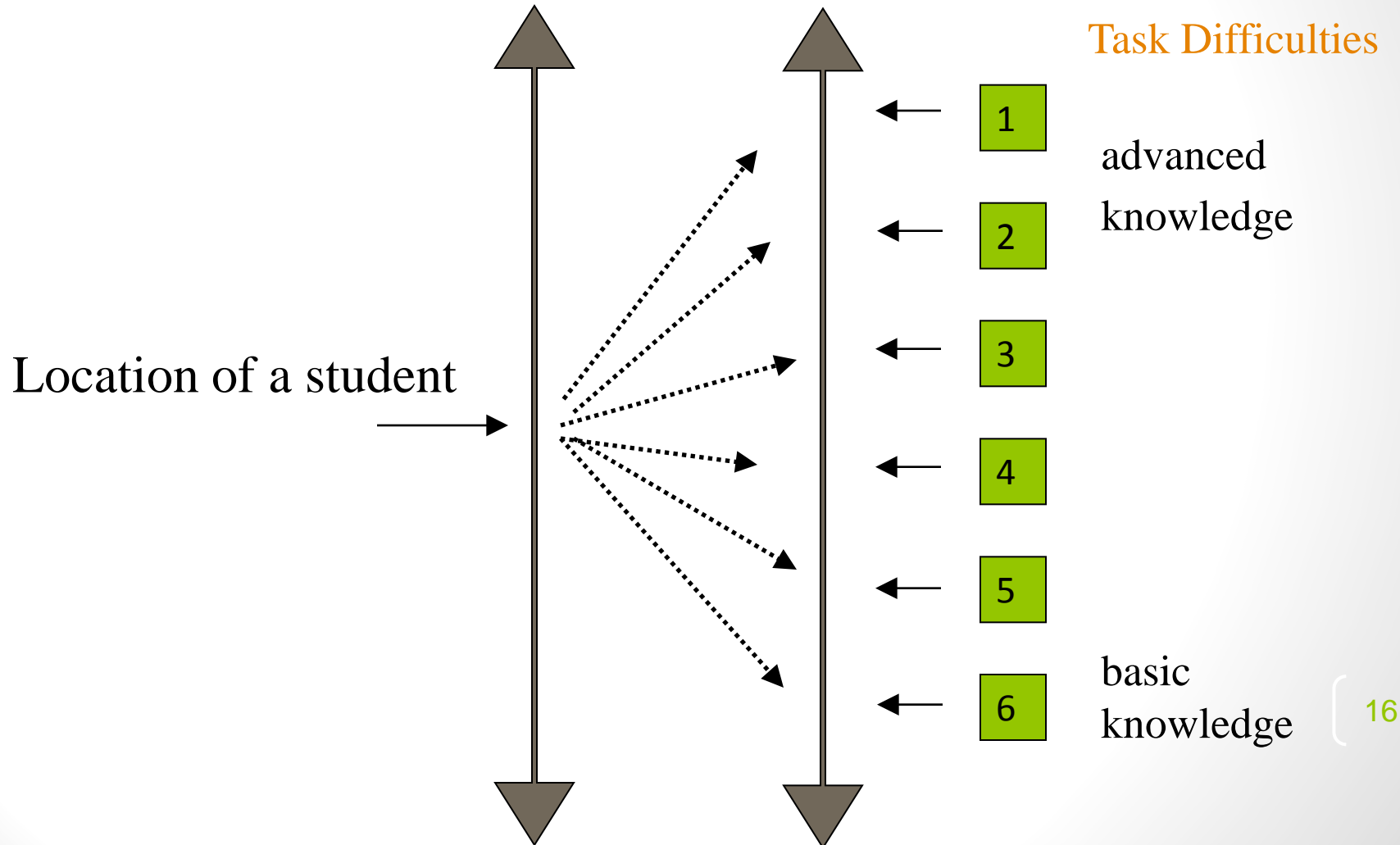
Variation in item difficulty



Variation in item difficulty

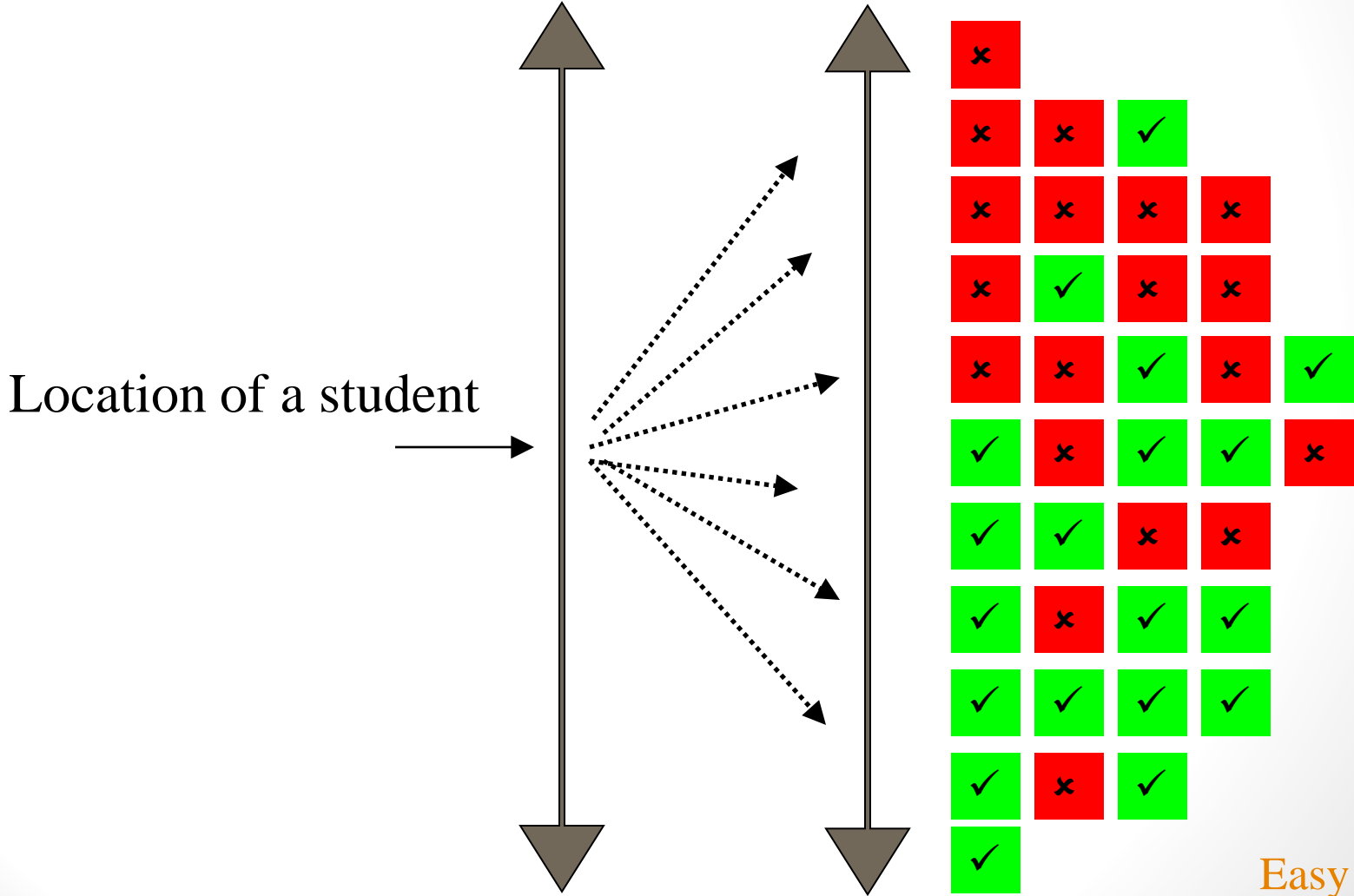


Comparing Students and Items



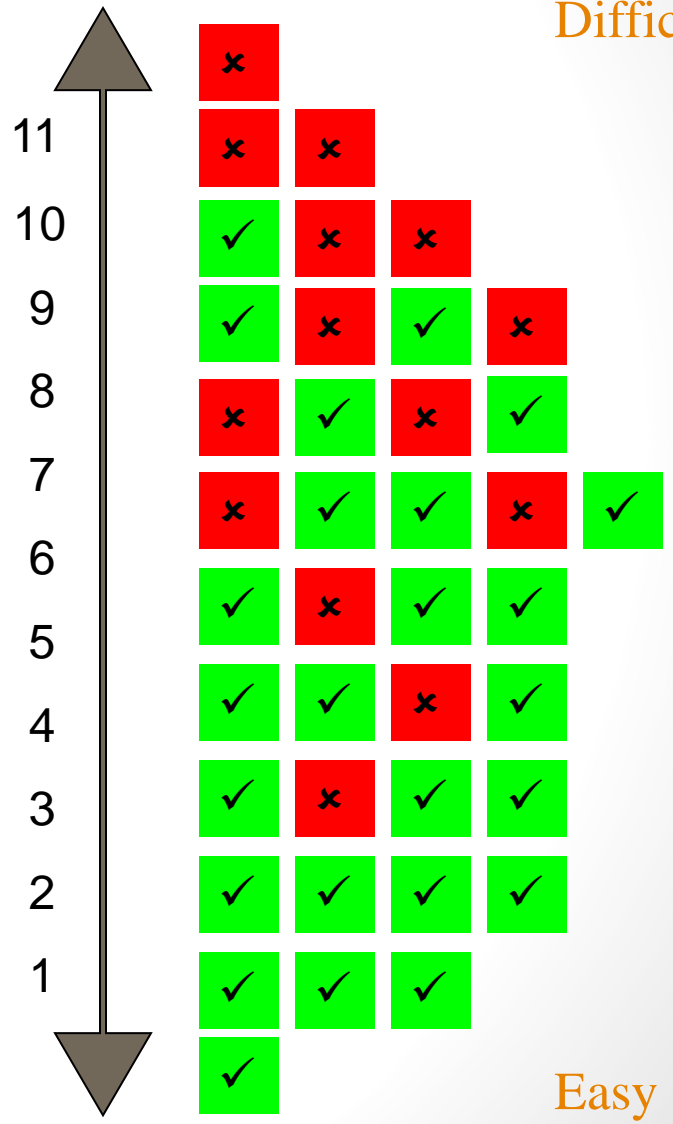
Comparing Students and Items

Difficult



Comparing Students and Items

Location of a student?
→

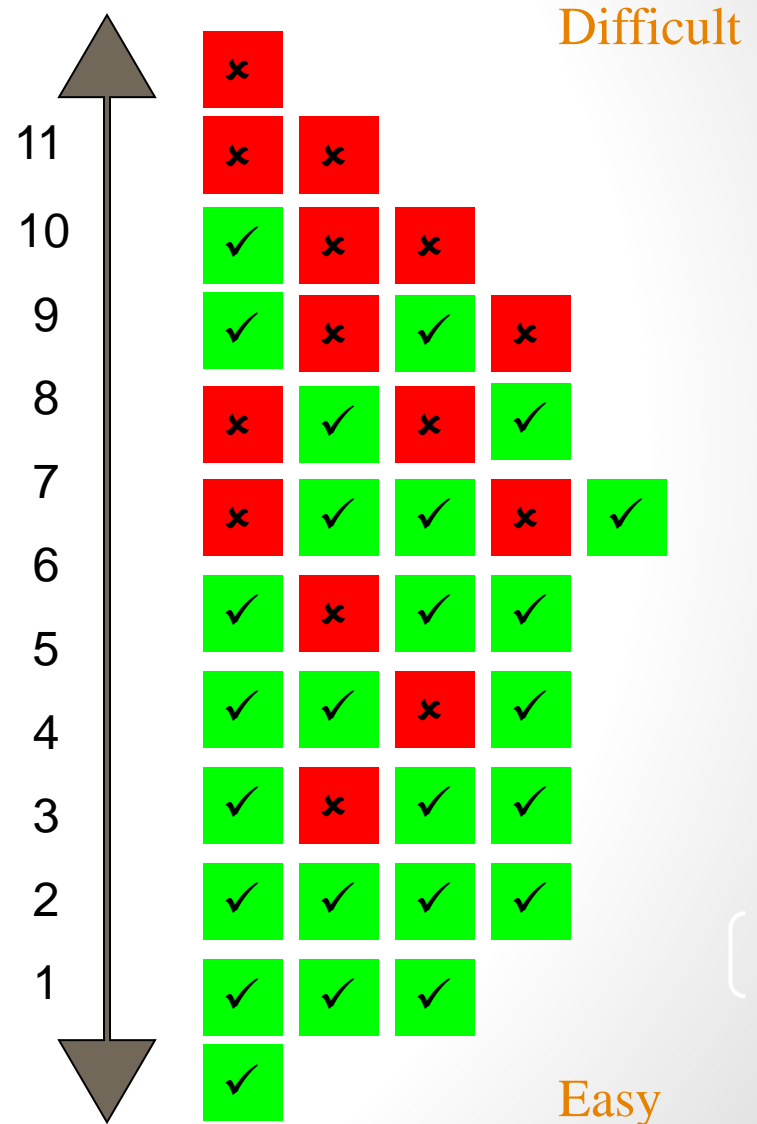


Difficult

Easy

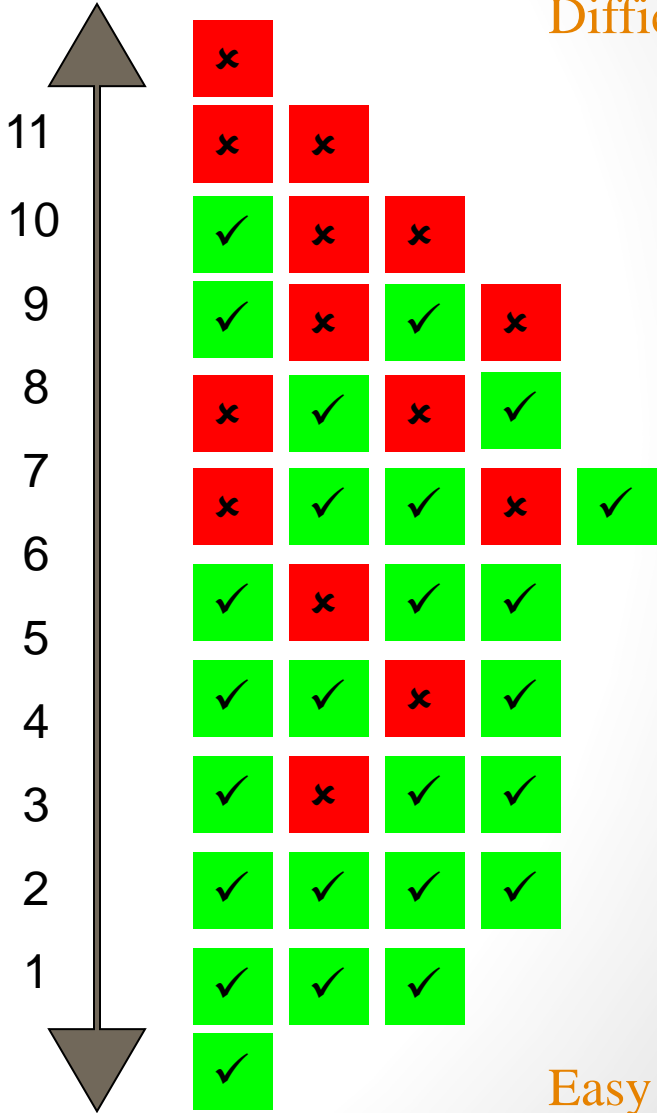
Comparing Students and Items

Location of a student?



Comparing Students and Items

Location of a student?

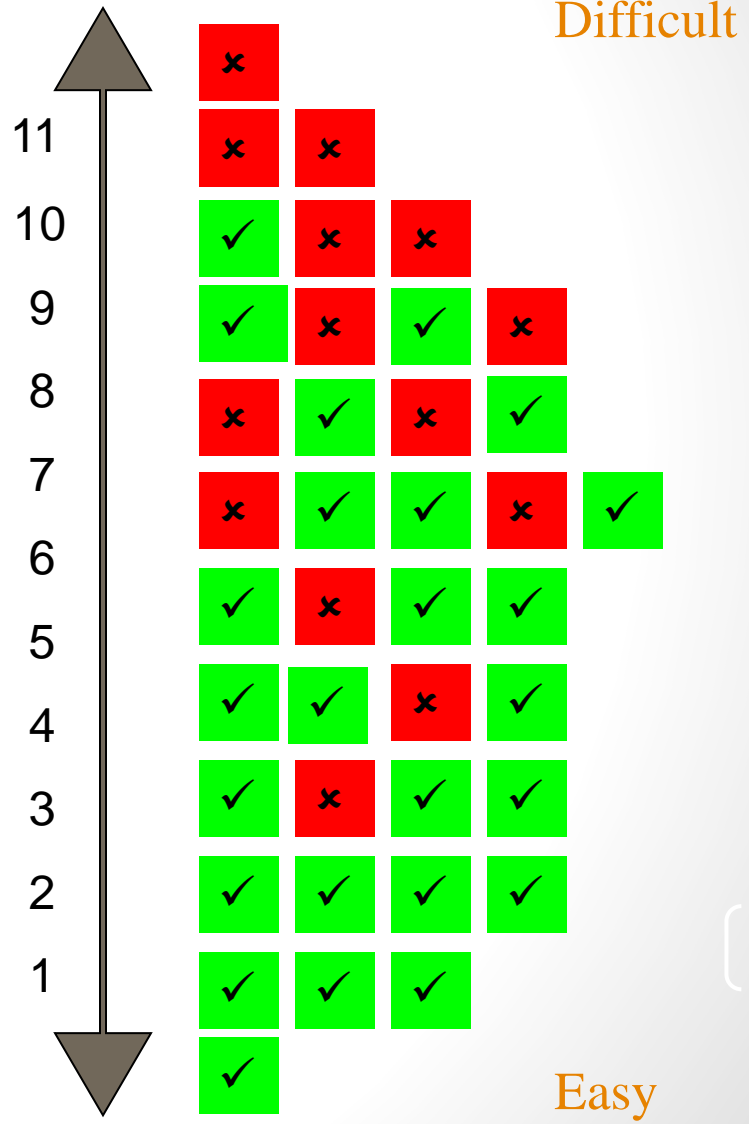
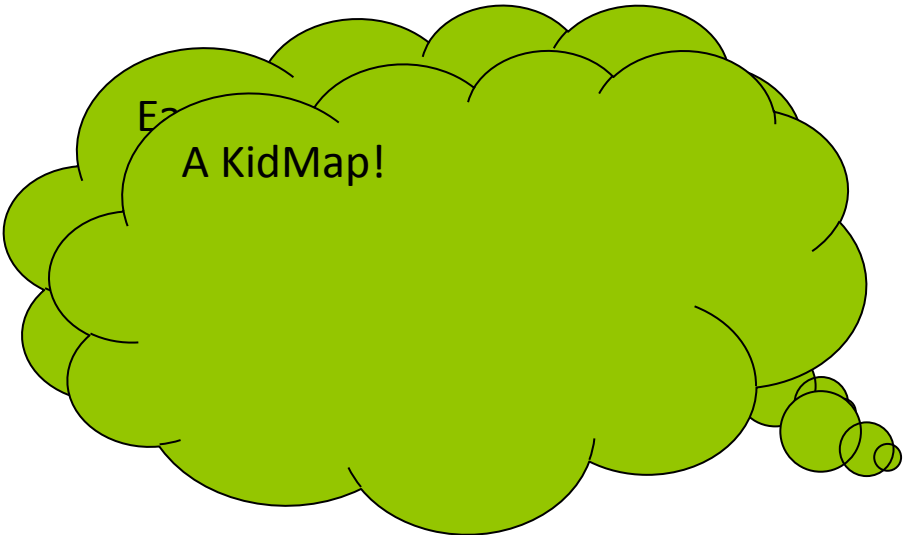


Difficult

Easy

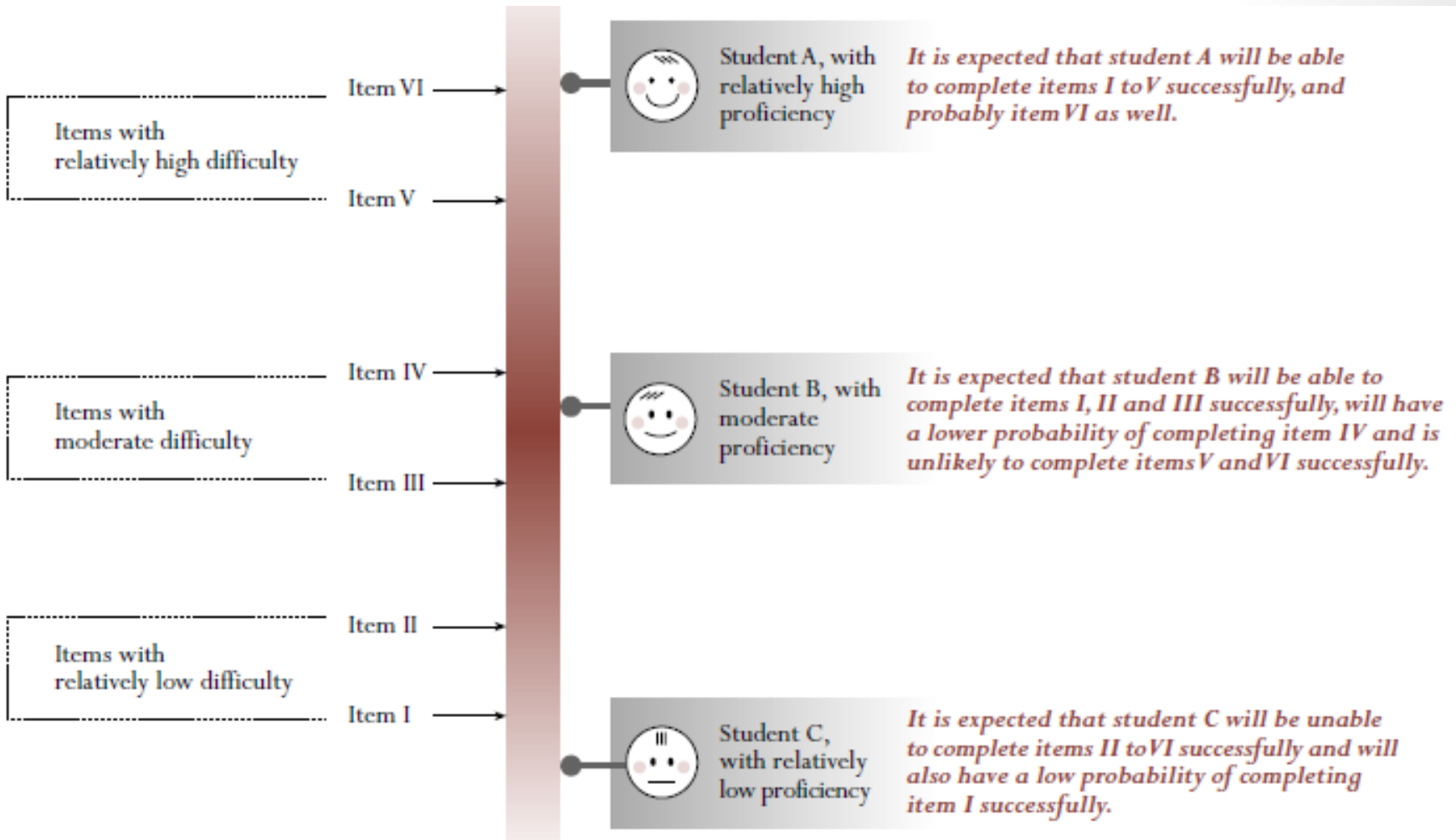
Sort correct and incorrect responses

Location of a student?



Constructing Proficiency Scales

- Step 1: carry out a skills audit of items
- Step 2: locate the skills along the ability scale
- Step 3: decide on band level cut-off values along the ability scale, and response probability.
- Step 4: write summary descriptions of the skills for each band level
- Step 5: calculate student abilities and place students in levels.
- Step 6: decide on any transformations of scaled scores.
- Step 7: compute cohort statistics such as percentages in levels.



4

At Level 4, students can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic ones, linking them directly to aspects of real-world situations. Students at this level can utilise well-developed skills and reason flexibly, with some insight, in these contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments and actions.

3

At Level 3, students can execute clearly described procedures, including those that require sequential decisions. They can select and apply simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They can develop short communications reporting their interpretations, results and reasoning.

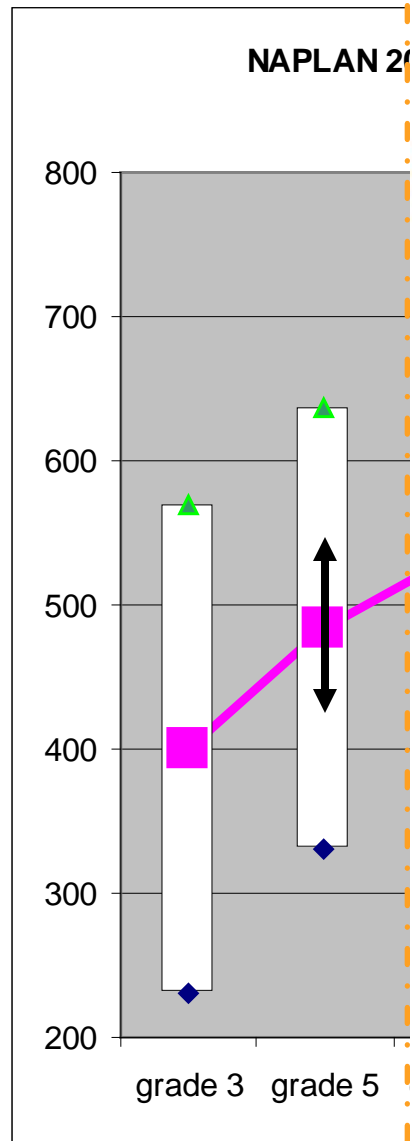
2

At Level 2, students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make

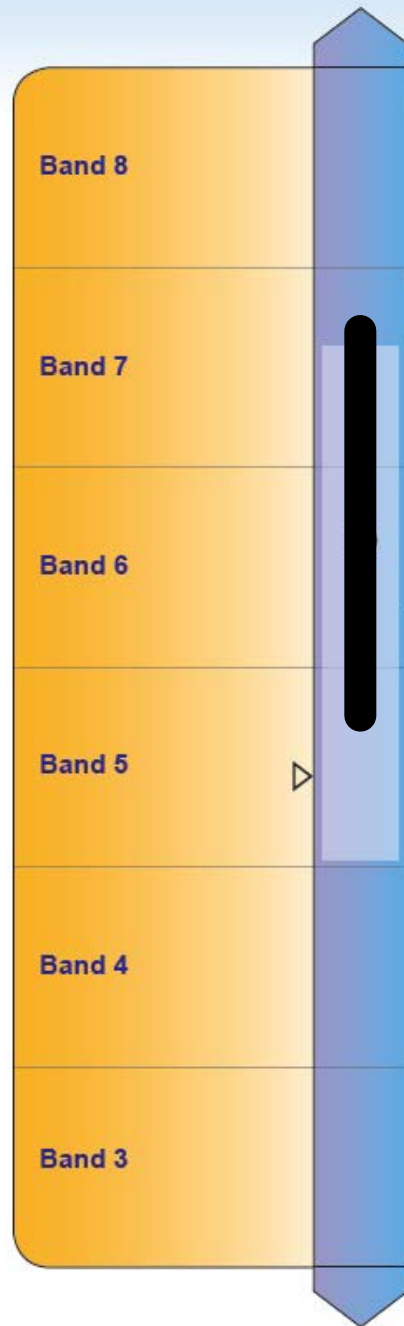
Limitations

- Descriptions are generic – They may not apply to individuals well
- Error margin of student score is typically not shown

On the NAPLAN



Year 5



Cautions about Reporting

Stephen Campbell

Year 4 Semester 1

Learning Area	Rating	Year 2	Year 3	Year 4	Year 5	Year 6
Health and Physical Education	E	●				
Interpersonal Development	B				●	
Personal Learning	B				●	
Civics and Citizenship	C			●		
The Arts	D		●			
English Reading	R					●