

# The role of assessment in evaluation

ERI-SEE Zagreb  
December 17, 18, 2009

Frans Kleintjes  
Cito, The Netherlands

# Eight steps of test construction

1. Purpose of the test
2. Test specification
3. Item construction
4. Pretesting
5. Item evaluation
6. Test assembly
7. Frame of reference
8. Manual and justification



# A mixed variety of purposes

Monitoring educational levels  
Placement  
Guaranteeing qualification

Allocation  
Prognosis

Selection  
Communication  
Achievement  
Operationalisation  
Evaluation

Instruction  
Steering the learning process  
Socialization  
Progress  
Qualification  
Diagnosis



now you know

# Purposes of educational tests

‘..... they **provide information to support decision making**. However, they differ in the kinds of information they seek and the types of decisions they can support and test results appropriate for some decisions may be inappropriate for other.’

(source: P.A. Barton, *Testing in American schools*, 1994)



now you know

# Quality of information

The *Standards for Educational and Psychological Testing* suggest that:

“ ...the validity of an intended interpretation of test scores relies on all the available evidence relevant to the technical quality of a testing system. This includes evidence of careful test construction ... “



now you know

# Test: definitions

- A test is an objective and standardized method for estimated behavior, based on a sample of that behavior.
- A standardized test is a test that uses uniform procedures for administration and scoring in order to assure that results from different people are comparable.



now you know

# Three basic functions of educational tests

1. Monitoring at system level
2. Monitoring at individual level
3. Assessment for individual decisions

# Monitoring at system level (1)

To judge the quality of an educational system.

- Macro level (national, international)
- Meso level (regional, school)



now you know

# Monitoring at system level (2)

## Functions

Used for monitoring and making administrative decisions about aggregated groups of students (e.g. a school, instructional programme, curriculum, district, province, nation)

## Examples

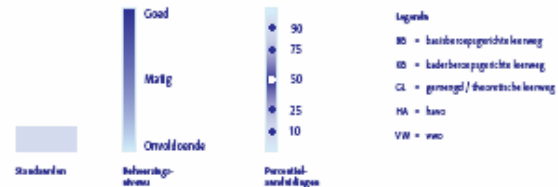
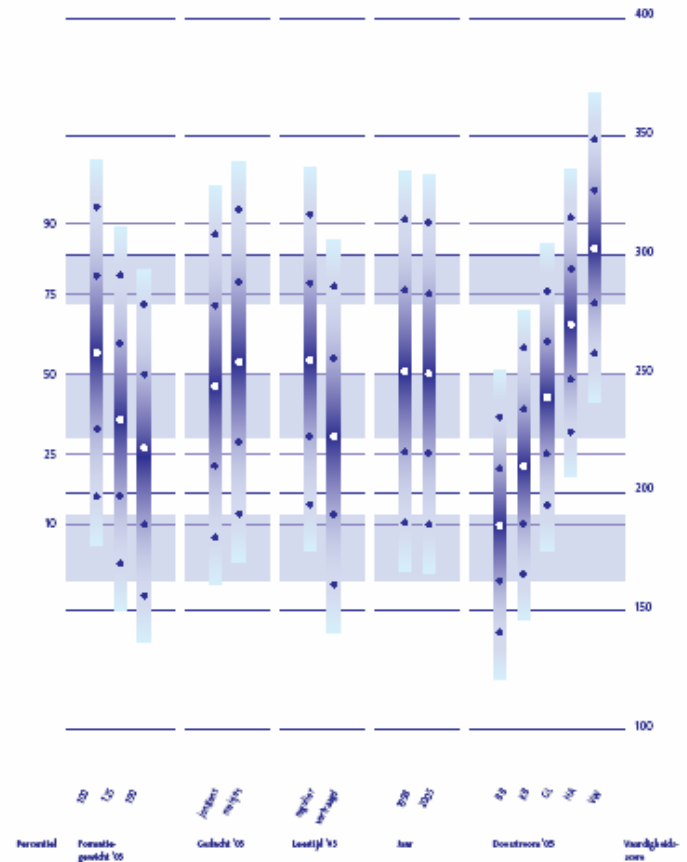
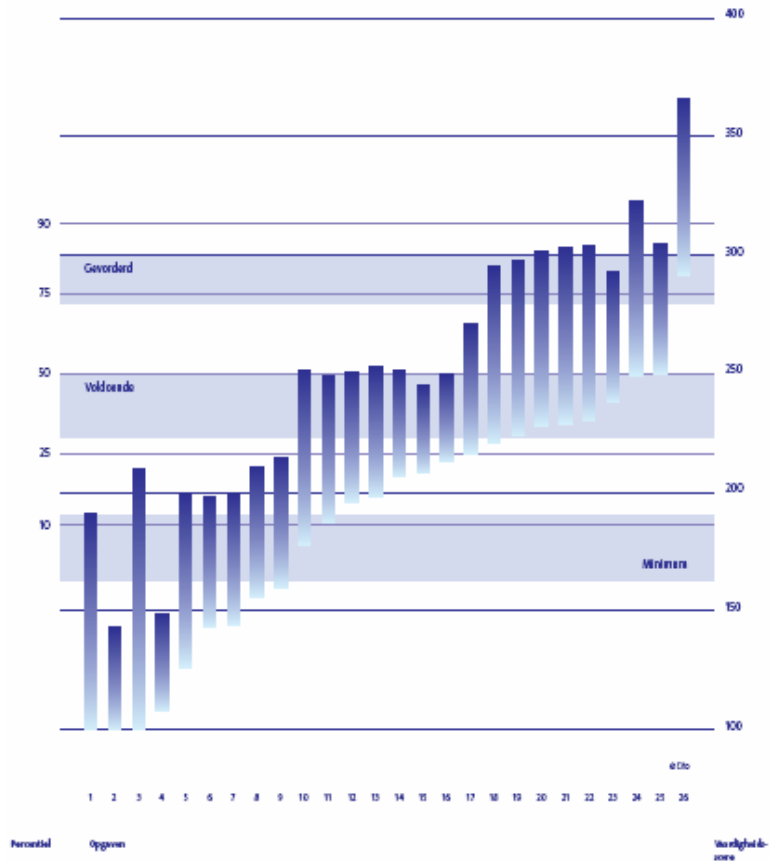
- Make a decision about instructional programmes and curriculum change
- Evaluate experimental and innovative programmes
- Provide general information about performance of the overall educational system
- Report to parent and school board about a school's or a districts performance



now you know

# Example of system evaluation

De vaardigheidschaal bij het onderwerp  
Begrijpen van geschreven teksten



# Example of school evaluation

# End of Primary School Test

## School Report A

# eindtoets basisonderwijs

School 67890  
CITOSCHOOL  
  
Schoolstraat 123  
1234 AB CITOCITY

Number of Pupils 2007: 52 of 52  
Show Selection list 2006: 24 of 24  
2005: 26 of 26  
  
Printdate 5-3-2007

### Explanation

Standard scores	Number questions	Average st. score	-2	-1	0	+1	+2
End of Primary School Test 2007	200	534,0			X-		
End of Primary School Test 2006	200	533,5			X---		
End of Primary School Test 2005	200	532,9			X----		
	Number questions	Average % of corr. answers					
<b>End of Primary School Test 2007</b>	<b>200</b>	<b>73%</b>			<b>X-</b>		
Explanation part 1 of the table							
			-2	-1	0	+1	+2
	Number questions	Average % of corr. answers					
<b>Language</b>	<b>100</b>	<b>72%</b>			<b>I---</b>		
1 Writing	30	66%			*I		
2 Spelling of verbs	10	65%			***I		
3 Spelling (all other than verbs)	10	49%			*****I		
4 Reading comprehension	30	87%			I*****		
5 Vocabulary	20	66%			I*****		
<b>Arithmetic/Mathematics</b>	<b>60</b>	<b>61%</b>			<b>I-----</b>		
6 Dealing with numbers and mental arithmetic	25	52%			*I		
7 Proportions, fractions and percentages	20	56%			*****I		
8 Geometry, money and time	15	58%			I*****		



# End of Primary School Test

## IC-standard scores and IC-percentiles

# eindtoets basisonderwijs

**School** 67890  
CITOSCHOOL

Schoolstraat 123  
1234 AB CITOCITY

**Number of Pupils** 23 of 23  
[Show Selection list](#)

**Printdate** 7-3-2007

### Average standard scores EPST

- 1. Uncorrected : 535,5
- 2. Corrected for APW : 534,7 interval 532,1 – 537,3
- 3. Corrected for APW and RC : 536,0 interval 534,6 – 537,4

	Percentiles corrected for Average Pupil Weight (APW)			Percentiles corrected for APW and Reading Comprehension (RC)		
	lower boundary	IC-Percentile	upper boundary	lower boundary	IC-Percentile	upper boundary
<b>Language</b>	18	50	82	54	88	99
<b>Producing texts</b>						
• Writing of texts	29	66	91	55	90	99
• Spelling of verbs	8	28	60	9	33	68
• Spelling (all other than verbs)	28	67	93	37	80	98
<b>Understanding Texts</b>						
• Reading Comprehension	5	27	66	n.v.t.	n.v.t.	n.v.t.
<b>Vocabulary</b>	28	63	90	44	81	97

# Monitoring at individual level (1)

To judge the quality of the educational learning process.

- Micro level (individual)
- Formative

# Monitoring at individual level (2)

## Functions

Used to monitor the learning process:

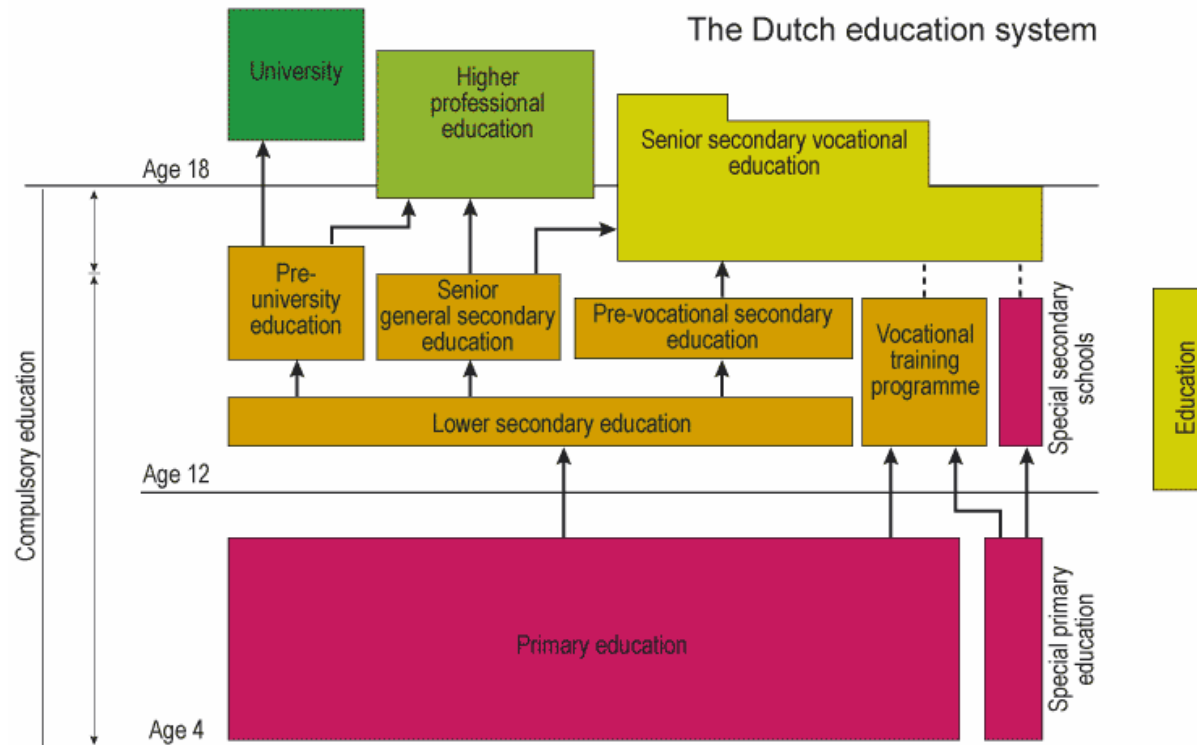
- progress of each student
- to inform student on the individual learning process
- to inform teacher on teaching decisions

## Examples

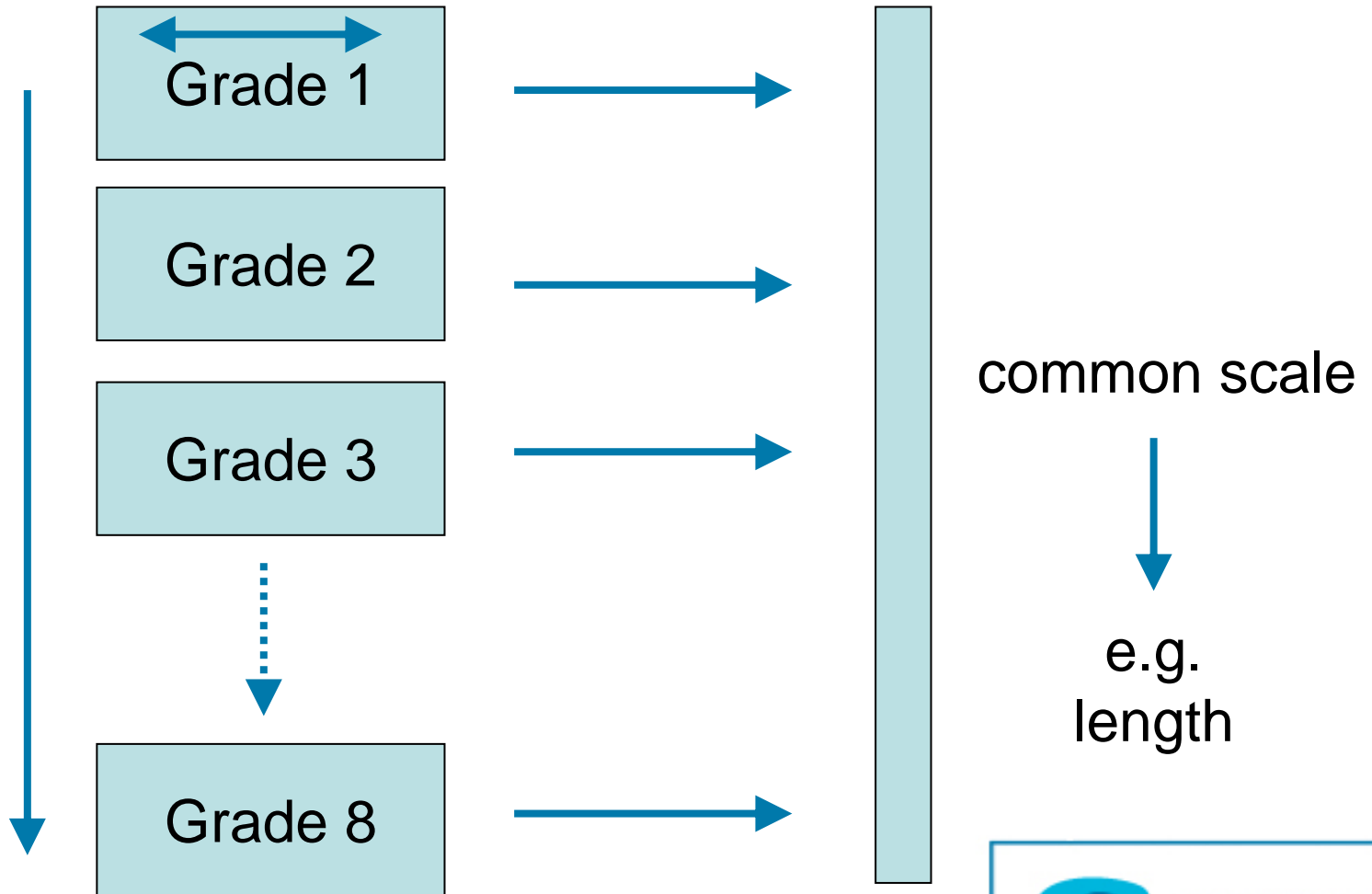
- Diagnose each student's strengths and weaknesses
- Monitor mastery and understanding of new material
- Adapt curriculum to progress as indicated by the test
- Monitor the effect of a lesson or unit
- Plan lessons that build on student's level of current understanding

# Monitoring at individual level (3)

## Example: Student monitoring system



# Monitoring at individual level (4)



# Monitoring at individual level (5)

Laura



116



Item:  
'new'

112

108



'blue'

104

100



'severe'

96

Peter<sup>2</sup>



92



'scale'

88

Peter<sup>1</sup>



84

80



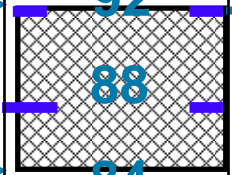
'cloud'

76

72



'button'



# Assessment for individual decisions (1)

To judge the quality or potential of a student.

- Prediction or classification
- Placement in special programmes
- Selection
- Certification

# Assessment for individual decisions (2)

## Functions

Used to allocate educational resources and opportunities among individuals

## Examples

### *Prediction or classification*

- Placement in school type most fit

### *Placement*

- Place student in special programme (remedial, gifted and talented)

### *Selection*

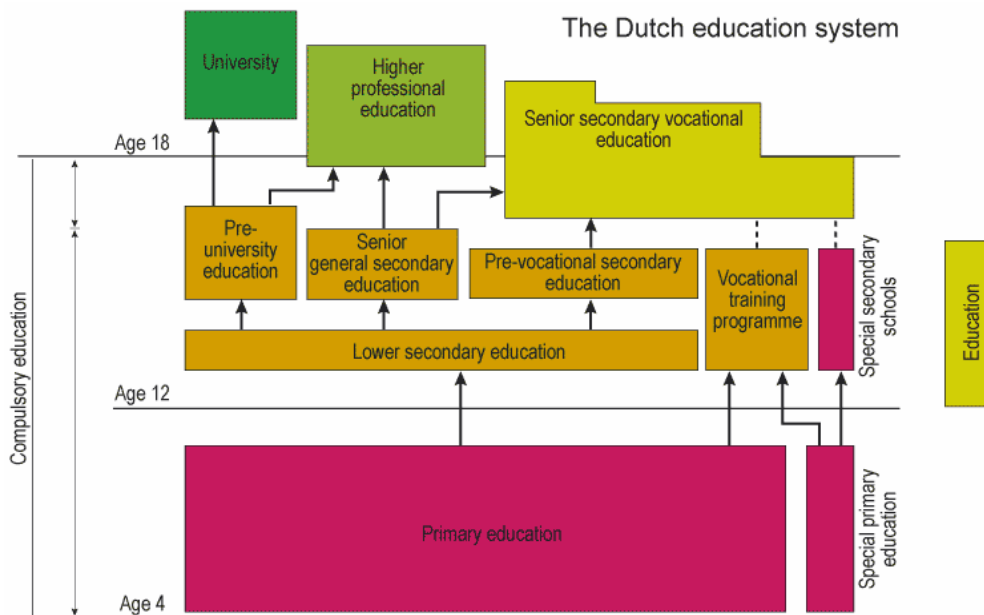
- Admission to schools or universities

### *Certification*

- Make a decision about grade promotion
- Certify mastery of a course or study
- School diploma

# Example of predictive test (1)

Advice in transmission from primary education to secondary education (classification)





# Framework for assessment

- **Level:**  
System (International/national) - School – Student  
Primary – Secondary - Vet - HE
- **Type:**  
Assessment *of / for* learning
- **Purpose**  
Intended purpose-Problems encountered -Impact
- **Responsibility**  
Who is initiator? Who benefits? Who pays?