

# Scientific Concepts Test: parallel semantically equal tests development in two languages



Tatjana Kanonire  
Malgožata Raščevska

01.04.2015

# Why do we need equal tests in Latvia?

---

- ❑ **Latvian** is the official state language, and **Russian** is the native language of a large minority within the community.
- ❑ The community supports separate schools for children whose native language is Latvian or Russian.
- ❑ The curriculum is the same for Latvian and Russian schools.
- ❑ The education curriculum in the Russian minority school system is bilingual, i.e. instruction is presented in the Russian language and the Latvian language is offered as a second language.
- ❑ Latvian is introduced progressively from primary school through the 9th grade, with the average proportion of instruction in the 9th grade being about 40% and 60% in Russian and in Latvian respectively.

# Verbal comprehension

---

- Verbal comprehension (VC) is the **ability to perceive and understand linguistic units** (Sternberg, & Powell, 1983).
- **The basis of VC is formed by concepts comprehension**, as concepts are the part of individual's mental representations (Rosch, & Mervis, 1975; Ross, 2000; Murphy, & Medin, 1985).
- **The ability to comprehend single concepts is very important for complex idea comprehension** which are expressed in written or spoken sentences (Swanson, Saez, Gerber, & Leafstedt, 2004; Berninger, Abbott, Swanson, et. al, 2010 ).

# Verbal comprehension (2)

---

- In psychology VC is studied by:
  - Vocabulary tasks,
  - Reading tasks,
  - Reasoning tasks with verbal and non-verbal stimuli,
  - Awareness tasks,
  - Tasks of situation comprehension.

We can distinguish two types of vocabulary tasks:

- Tasks for vocabulary breadth,
- Tasks for vocabulary depth.

# Verbal comprehension (3)

---

- An example of vocabulary depth task:
  - We ask person to give the definition of particular words

What is TABLE?

Possible answers:

- It is furniture,
  - We use it for work and dinner,
  - It has four legs and large board on the top,
  - I have it at home.
- So, we receive information about both:
    - vocabulary – does the person know the meaning of the concept,
    - verbal abilities – abstract or concrete thinking.

# Verbal comprehension (4)

---

=> assessment of scientific concepts vocabulary is appropriate for the evaluation of student comprehension of scientific concepts.

Concept defining task is more appropriate for scientific concepts comprehension assessment among adolescents.

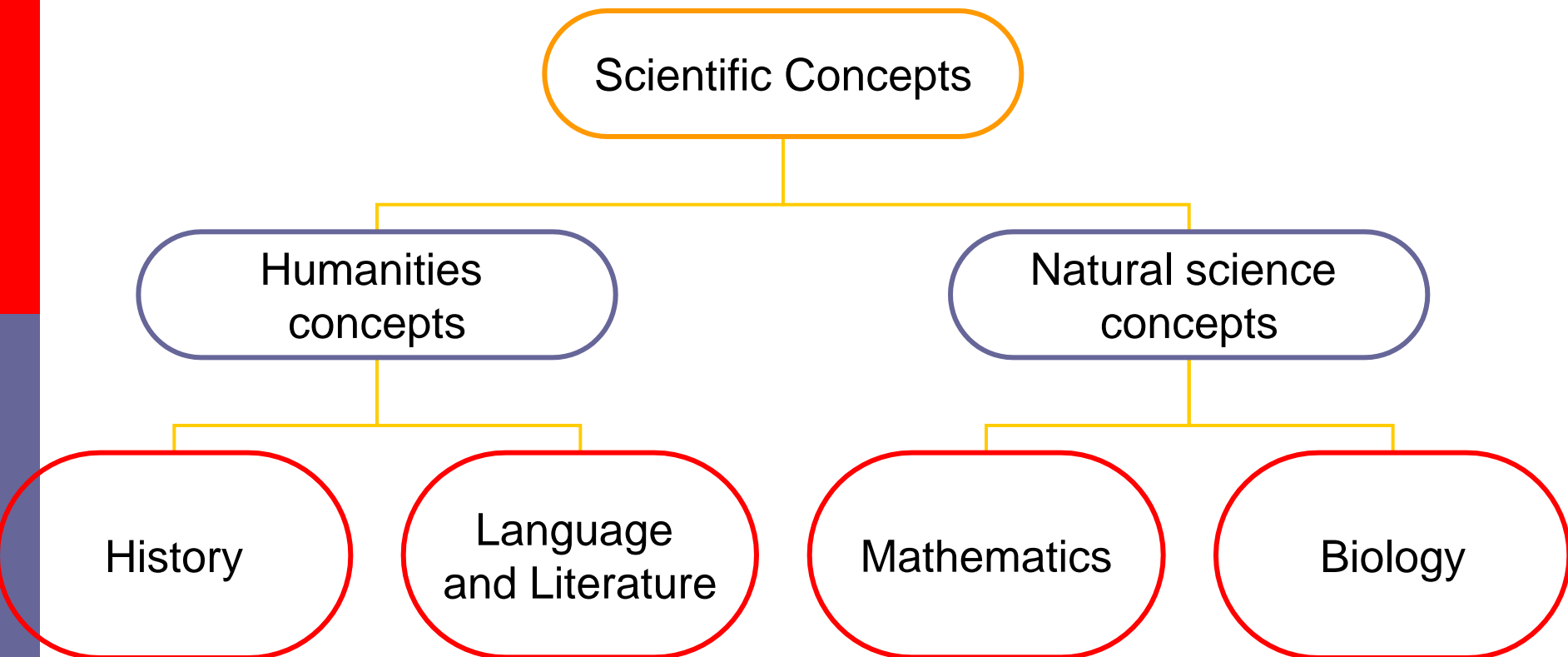
# Can we adapt a test?

---

- Differences in educational curriculum;
- Two test types are commonly used:
  - final examination tests - focuses on the resultant acquired knowledge of subject content that has been studied,
  - tests that focus on the general skills which students need for studying (i.e. Woodcock-Johnson III Test of Achievement - Woodcock, McGrew, & Mather, 2001).
- ⇒ neither of these tests types includes a vocabulary component, ***vocabulary is integrated into the test tasks in ways that prevent their direct measurement.***
- Trying to adapt a test with verbal stimuli in two languages can cause some biases.

# Tests development

- Target population is Latvian and Russian native speakers from the 9th grade (aged 14-16).
- The content is based on Latvian Primary Education Standard for the 9th grade (LR Izglītības Ministrija, 2008).





# Tests development (2)

---

- The initial item pool of 52 items was compiled on the basis of expert recommendations from the 9th grade teachers of Latvian and Russian students in the subject areas of history, literature, native language, mathematics and biology.
- Items were retained in the tests that appeared in both languages and were taken from the expert listings that matched true criteria.

# Tests development (3)

---

Items in both languages:

- are semantically similar, e.g. “segment” has the same meaning in both languages
- have a different phonology  
“рабство” – “verdžība” (“slavery”)  
NOT “жанр” - “žanrs” (“genre”)
- are represented in one word  
NOT “лирический герой”

# The aim of the study

---

- The research examined the extent to which psychometric properties of both the SCT-L and the SCT-R met validity and reliability standards (Standards of Educational and Psychological Testing, 1999; Kline, 2000).

# Method

---

## Participants

The **Latvian sample** consisted of 175 students with native Latvian aged between 14 and 16 ( $M = 15.10$ ,  $SD = 0.50$ ; 49% - boys)

**Subsamples from Latvian sample** were used for verifying SCT-L validity:

- (a) 22 respondents aged between 14 and 16 ( $M=15.50$ ,  $SD = 0.96$ , 50% boys) completed WISC-IV Vocabulary subtest,
- b) 66 respondents aged between 14 and 16 ( $M = 15.06$ ,  $SD = .49$ , 49% boys) completed the Latvian Vocabulary test (LVT).

# Method (2)

---

## **Participants**

The **Russian sample** consisted of 101 students with native Russian language aged between 14 and 16 ( $M = 15.12$ ,  $SD = 0.48$ ; 51 % were boys).

**Subsamples from Russian sample** were used for verifying SCT-R validity:

- 51 respondents aged between 14 and 16 ( $M = 15.12$ ,  $SD = 0.52$ , 55 % boys) completed the Russian Vocabulary test (RVT).

All participants were students of the 9th grade.

# Method (3)

---

## Instruments

### **Latvian and Russian versions of Scientific Concepts Test (SCT-L and SCT-R)**

- The initial item pool of each test consisted of 52 items. The number of items was eventually reduced to 26 based on item data.
- The following scoring system was used in the tests: 2 points were given if the answer was correct and the key words were mentioned; 1 point was given if only one key word was mentioned, a concrete example was given, or if the answer was partly correct; No points were given if the answer was incorrect.

# Examples

---

## □ “Sekante” – “Секущая” (“Secant line”)

Definition: a straight line that intersects the circle at two points.

### 2-point answers:

- прямая, проходящая через окружность,
- линия, проходящая через окружность и касающаяся ее в двух точках

### 1-point answers:

- прямая, проходящая через две точки фигуры
- та, которая идет через что-то, проходит через круг

### 0-point answers:

- пересекает фигуру,
- отрезок, который проходит насквозь линию, угол геометрической фигуры,
- отсекающая линия

# Examples (2)

---

- “Šūna” – “Клетка” (“Cell”)
- Definition: the smallest living system, in which the main characteristics of life are presented.

## **2-point answers:**

- микроскопический организм, который является составляющей любого сложного организма,
- это мельчайший организм, без нее не возможна жизнь

## **1-point answers:**

- бывает кровяной
- микроскопическая часть тела из которой сделана наша кожа
- микроорганизм

## **0-point answers:**

- то, что находится под кожей у человека, животного,
- одна из тканей живых существ,
- маленький микроб



# Method (4)

---

## Instruments

- **Latvian Vocabulary Test (LVT) and Russian Vocabulary Test (RVT)**
  - measure verbal comprehension. Each test consisted of 19 items and included widely used words taken from everyday life (e.g. “to wash”, “book”) with the task to define the words. Both the Latvian and Russian versions contain identical concepts.
  - The scoring system as in SCT is used in the tests. The Spearman-Brown split-test reliability index of the LVT was .73; the Spearman-Brown split-test reliability index of the RVT was .68.

# Method (5)

---

## **Instruments**

**The WISC-IV Vocabulary subtest Latvian version** (Wechsler, 2003; Raščevska, Sebre, & Ozola, 2011, September) was used in order to verify the validity of the SCT-L. WISC-IV item assessment principles were applied to evaluate the items in the Latvian version.

# Method (6)

---

## Procedure

Both Latvian and Russian groups completed the tests either in writing or by oral response.

Writing response form was used in the pilot study and so as not found statistically significant differences, it was decided to include respondents who completed test in writing into the analysis.

- Among the Latvian sample 13 % of participants ( $M = 22.96$ ,  $SD = 7.84$ ) completed the test in writing while 87% completed test orally ( $M = 22.41$ ,  $SD = 8.77$ ). There were no significant differences between participants who completed test in writing or orally ( $t(176) = .28$ ,  $p > .05$ ).
- Among the Russian sample, 33 % of participants ( $M = 25.47$ ,  $SD = 9.95$ ) completed the test in writing and 67% completed the test orally ( $M = 25.15$ ,  $SD = 8.58$ ). There were also no significant differences between participants who completed test in writing or orally ( $t(99) = .17$ ,  $p > .05$ ).

# Methods (7)

---

## Procedure

- Data collection for both language versions of the tests was undertaken simultaneously.
- Two experts undertook the evaluation of participant SCT-L, SCT-R, LVT and RVT responses.
  - At first, a common evaluation procedure was developed and then each expert independently evaluated the data. Then these evaluations were compared and any difficult issues were discussed and resolved. Identical evaluation criteria were applied for both the Latvian and Russian versions of the test. Identical answers in Latvian and Russian versions were scored similarly.

# Results

---

Several concepts were omitted because the respondents indicated:

- a different semantic understanding of these concepts
  - e.g. "cunfte" ("цех") stands for "craftsmen's union" in Latvian; however, in Russian, this concept has several meanings: "craftsmen's union", "place where something is manufactured", with the most common being the second meaning);
  - e.g. "temple" ("храм") in Latvian concern only to non-Christian religions, usually to paganism.
  
- there were some differences in the difficulty of explaining the various concept (e.g. for some items a synonym was easily to found in one language and not in another).

**Table 1a. Item difficulty and discrimination indices of SCT-L and SCT-R**

Scale and item	SCT-L			SCT-R		
	M	SD	Discrimination index*	M	SD	Discrimination index*
<b>Natural science concepts scale (total)</b>	<i>12.63</i>	<i>5.44</i>		<i>13.89</i>	<i>5.35</i>	
Segment ( <i>отрезок</i> )	1.46	.80	.43	1.38	.83	.33
Stomach ( <i>желудок</i> )	1.36	.71	.50	1.50	.66	.24
Triangle ( <i>треугольник</i> )	1.29	.81	.42	1.74	.56	.32
Joint ( <i>сустав</i> )	1.11	.75	.45	.93	.86	.54
Reproduction ( <i>размножение</i> )	1.09	.85	.38	1.30	.82	.50
Development ( <i>развитие</i> )	.94	.52	.30	.88	.67	.42
Cell ( <i>клетка</i> )	.89	.78	.40	.93	.79	.62
Breathing ( <i>дыхание</i> )	.87	.70	.36	1.01	.75	.36
Area ( <i>площадь</i> )	.77	.90	.36	.67	.66	.55
Tangent line ( <i>касательная</i> )	.74	.91	.52	.66	.83	.39
Plant ( <i>растение</i> )	.68	.62	.38	1.02	.69	.31
Angle ( <i>угол</i> )	.67	.74	.47	.70	.61	.26
Secant line ( <i>секущая</i> )	.41	.76	.33	.63	.83	.47
Equation ( <i>уравнение</i> )	.37	.58	.37	.52	.58	.47
<b>Item mean</b>	<b>.90</b>	<b>.74</b>	<b>.41</b>	<b>.99</b>	<b>.73</b>	<b>.41</b>

**Table 1b. Item difficulty and discrimination indices of SCT-L and SCT-R**

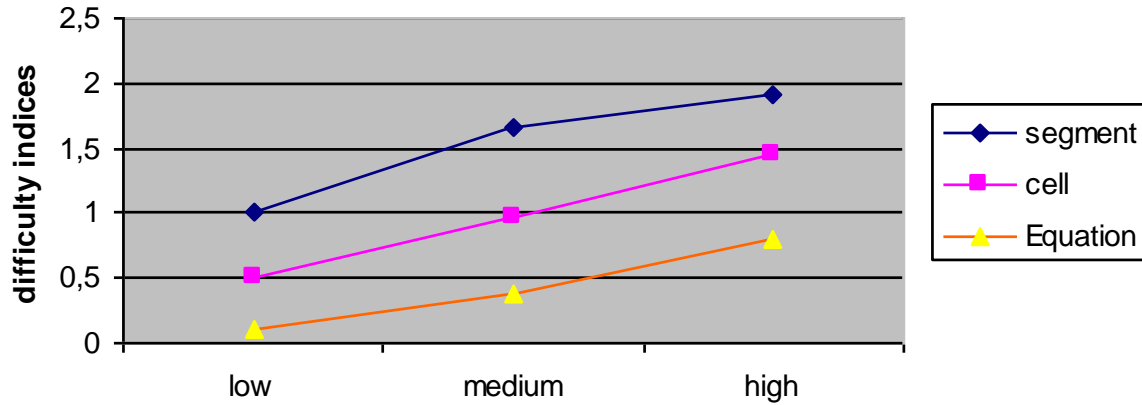
Scale and item	SCT-L			SCT-R		
	M	SD	Discrimination index*	M	SD	Discrimination index*
<b>Humanities concepts scale (total)</b>	9.85	3.82		11.27	4.27	
Monument ( <i>памятник</i> )	1.66	.54	.24	1.54	.59	.24
Society ( <i>общество</i> )	1.35	.68	.34	1.04	.47	.23
War ( <i>война</i> )	.99	.56	.38	1.16	.52	.41
Sentence ( <i>предложение</i> )	.99	.69	.38	1.14	.69	.43
Description ( <i>описание</i> )	.99	.76	.46	.98	.77	.30
Slavery ( <i>рабство</i> )	.92	.55	.35	1.34	.52	.20
Story ( <i>рассказ</i> )	.88	.64	.37	1.00	.68	.45
Pronoun ( <i>местоимение</i> )	.59	.67	.26	.87	.83	.37
Personification ( <i>олицетворение</i> )	.56	.76	.32	.61	.73	.33
Serfdom ( <i>крепостничество</i> )	.45	.74	.43	.59	.85	.54
Priesthood ( <i>духовенство</i> )	.30	.59	.24	.23	.56	.35
Nobility ( <i>дворянство</i> )	.16	.45	.27	.84	.85	.47
<b>Item mean</b>	<b>.82</b>	<b>.64</b>	<b>.34</b>	<b>.95</b>	<b>.66</b>	<b>.36</b>

**Table 2a. Item difficulty indices for respondents with different level of total score in Natural science concepts scale**

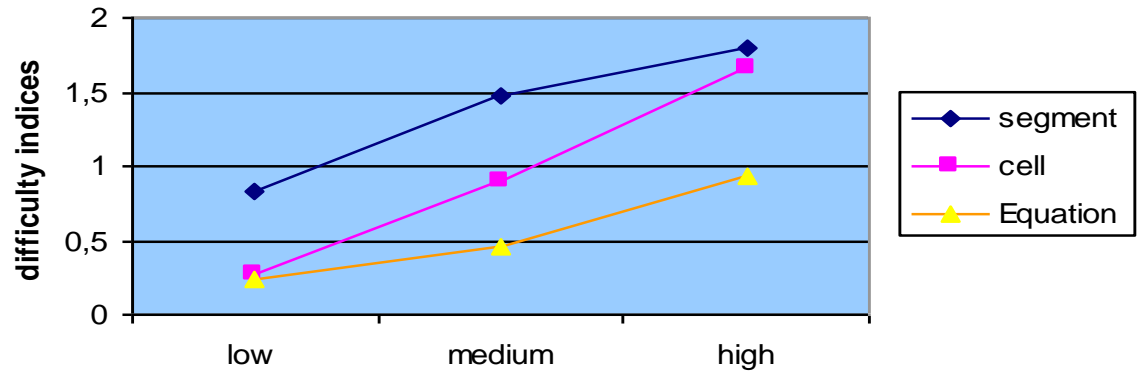
Scale and item	SCT-L			SCT-R		
	M1	M2	M3	M1	M2	M3
<b>Natural science concepts scale (total)</b>						
Segment ( <i>отрезок</i> )	1.00	1.65	1.92	0.83	1.48	1.79
Stomach ( <i>желудок</i> )	0.96	1.45	1.92	1.23	1.48	1.83
Triangle ( <i>треугольник</i> )	0.81	1.55	1.69	1.43	1.81	1.97
Joint ( <i>сустав</i> )	0.60	1.23	1.72	0.40	0.76	1.72
Reproduction ( <i>размножение</i> )	0.71	1.23	1.59	0.57	1.48	1.79
Development ( <i>развитие</i> )	0.39	0.78	1.03	0.33	0.98	1.31
Cell ( <i>клетка</i> )	0.51	0.97	1.44	0.27	0.90	1.66
Breathing ( <i>дыхание</i> )	0.69	1.03	1.23	0.63	0.95	1.48
Area ( <i>площадь</i> )	0.30	0.86	1.46	0.20	0.62	1.24
Tangent line ( <i>касательная</i> )	0.16	0.81	1.64	0.23	0.60	1.21
Plant ( <i>растение</i> )	0.54	0.93	1.36	0.73	0.95	1.41
Angle ( <i>угол</i> )	0.24	0.74	1.31	0.40	0.74	0.97
Secant line ( <i>секущая</i> )	0.07	0.43	0.97	0.17	0.48	1.34
Equation ( <i>уравнение</i> )	0.11	0.38	0.79	0.23	0.45	0.93



**The items difficulty indices of participants with different Natural science concepts total score (Latvian version)**



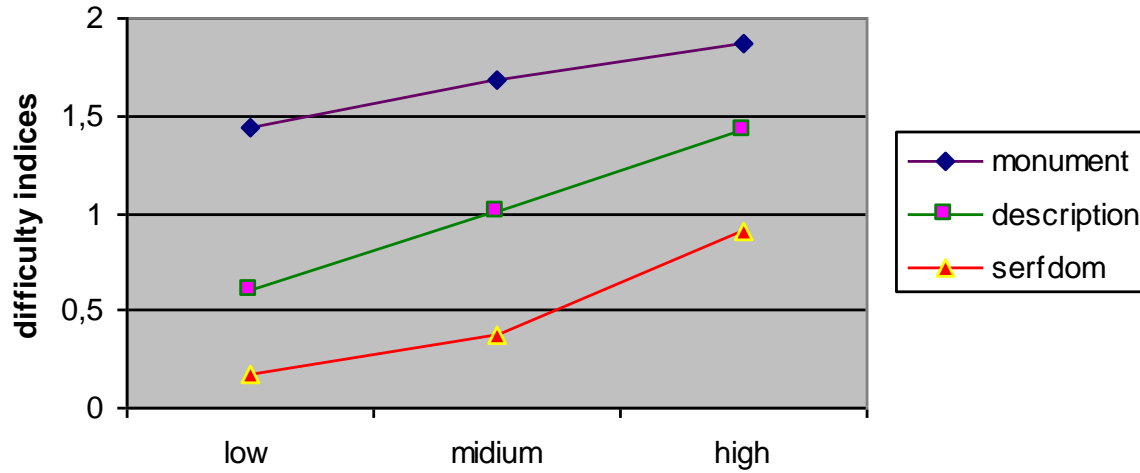
**The items difficulty indices of participants with different Natural science concepts total score (Russian version)**



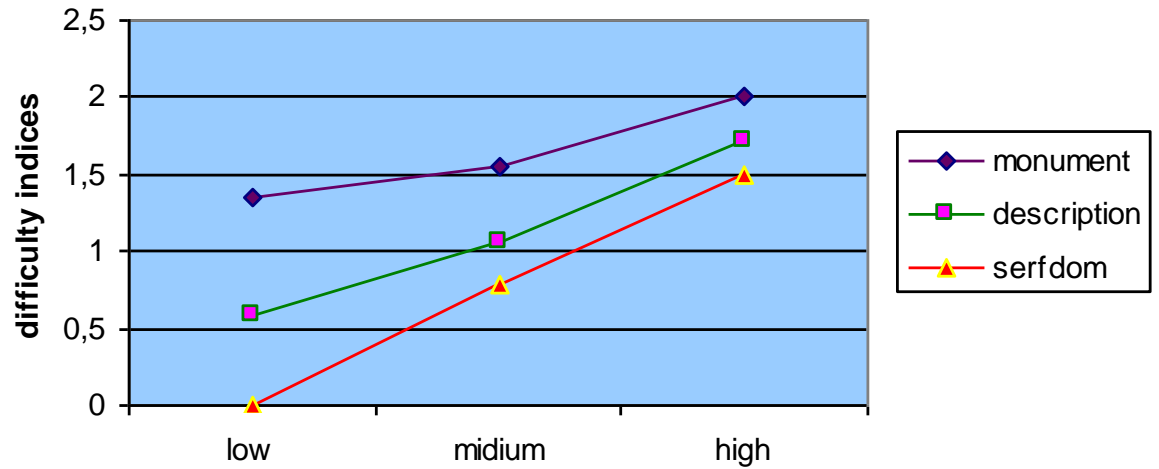
**Table 2b. Item difficulty indices for respondents with different level of total score in Humanities concepts scale**

Scale and item	SCT-L			SCT-R		
	M1	M2	M3	M1	M2	M3
<b>Humanities concepts scale (total)</b>						
Monument ( <i>памятник</i> )	1.44	1.69	1.87	1.35	1.56	2.00
Society ( <i>общество</i> )	0.67	0.95	1.16	0.86	1.14	1.14
War ( <i>война</i> )	0.77	0.99	1.27	0.97	1.14	1.71
Sentence ( <i>предложение</i> )	0.27	0.60	0.93	0.59	1.38	1.71
Description ( <i>описание</i> )	0.60	1.01	1.42	0.59	1.06	1.71
Slavery ( <i>рабство</i> )	1.10	1.37	1.60	1.22	1.34	1.71
Story ( <i>рассказ</i> )	0.10	0.51	1.18	0.70	1.04	1.64
Pronoun ( <i>местоимение</i> )	0.50	1.07	1.40	0.41	1.06	1.43
Personification ( <i>олицетворение</i> )	0.52	0.91	1.22	0.05	0.72	1.71
Serfdom ( <i>крепостничество</i> )	0.17	0.37	0.91	0.00	0.78	1.50
Priesthood ( <i>духовенство</i> )	0.12	0.28	0.56	0.05	0.22	0.71
Nobility ( <i>дворянство</i> )	0.10	0.07	0.40	0.27	1.08	1.50

The items difficulty indices of participants with different Humanities concepts total score (Latvian version)



The items difficulty indices of participants with different Humanities concepts total score (Russian version)



# The structure of the test

---

**The SCT-L and the SCT-R each have two scales:**

- Humanities Concepts Scale
  - Natural Science Concepts Scale,
- > that are established and based on the content validity, not on factorial validity.**
- Total score for SCT-L and SCT-R established in factor analysis
    - The principal components analysis of the SCT-L and of the SCT-R was done at the total score level, which revealed that the Humanities Concepts and Natural Science Concepts scales formed one component, for SCT-L explaining 87 % of the total item variance, with each scale eigenvalues at .93. (for SCT-R explaining 88 % of the total item variance, with each scale eigenvalues at .94).
    - Determinant, KMO and Bartlett test satisfy the criteria.

# Reliability

Table 3. Spearman-Brown split-test reliability indices for SCT-L and SCT-R

Scale	SCT-L	SCT-R
SCT Humanities Concepts scale	.73	.82
SCT Natural Science Concepts scale	.81	.81
SCT total score	.83	.86

# Concurrent and Convergent Validity (2)

---

- The SCT-L scales showed a significant positive correlation with the **WISC-IV Vocabulary** with the **Humanities Concepts Scale** ( $r = .58$ ;  $p < .01$ ) and with the **Natural Science Concepts Scale** ( $r = .51$ ;  $p < .05$ ) as well as with the **SCT-L total score** ( $r = .71$ ;  $p < .01$ ).
- The SCT-L also showed a significant positive correlation with the **LVT** with the **Humanities Concepts Scale** ( $r = .61$ ;  $p < .01$ ) and with the **Natural Science Concepts scale** ( $r = .73$ ;  $p < .01$ ) as well as with the **SCT-L total score** ( $r = .71$ ;  $p < .01$ ).
- The SCT-L showed a significant positive correlation with the **RVT** with the **Humanities Concepts Scale** ( $r = .53$ ;  $p < .01$ ) and with the **Natural Science Concepts Scale** ( $r = .48$ ;  $p < .05$ ) as well as with the **SCT-R total score** ( $r = .53$ ;  $p < .01$ ).

# Discussion

---

- Mean difficulty values for SCT-L items were lower than the mean difficulty values for SCT-R items, thus suggesting that the Latvian version of the test was more difficult than the Russian version.
- => One of explanations could be the difference in the level of knowledge of students.

# Discussion (2)

---

- The Spearman-Brown split-half test reliability indices for the SCT-L and the SCT-R scales satisfied the psychometric criteria.
  - But reliability estimates for the SCT-L Humanities Concepts Science Scale was only .73. This may occur because:
    - no items with difficulty indices in range between 1.35 to .99 (between the items “society” and “war”);
    - three items has an equal difficulty index of .99 (“war”, “sentence”, “description”)
- => items were retained because (a) the items represent different scientific disciplines, e.g. “war” – History, “sentence” and “description” – Language/Literature; (b) since the SCT-L and the SCT-R were developed simultaneously, one could not omit certain items from one version and leave them in the other. Similar problems were not seen in the SCT-R Humanities Concepts Scale.



# Discussion (3)

---

Further development of the SCT-L and the SCT-R:

- ❑ the collection of data on a broader representative sample, including samples matched for city and rural areas, as well as to sample a broader respondent age range is warranted.
- ❑ analysis of possible gender effect.
- ❑ Although the addition of items in each scale may be helpful, this goal may be difficult to achieve because of the limited amount of items which are considered valid for the scale itself and which are sufficiently phonologically different in the Latvian and Russian languages.
- ❑ verify the test-retest reliability.
- ❑ examine the issues of predictive validity (i.e. employing information from the school achievement index for the 9th grade), divergent validity and clinical validity.

---

Thank you for your attention!