Laboratory of Thinking – diagnosis of teaching science subjects in Poland – physics perspective

T GRECZYŁO1,2, D SOKOŁOWSKA3 and L DERYŁO

1 Institute of Experimental Physics, University of Wrocław, 50-137 Wrocław, pl. Uniwersytecki 1, Poland
2 Educational Research Institute, 01-180 Warszawa, ul. Górczewska 8, Poland
3 Smoluchowski Institute of Physics, Jagiellonian University, 30-348 Krakow, ul. Łojasiewicza 11, Poland

Abstract. The paper presents and discusses selected results of national diagnosis of teaching science subjects – physics, biology, chemistry and geography in Poland. The research was carried out at the beginning of the 2017/2018 school year and was the next instalment of research initiated in 2011 by Educational Research Institute (IBE) in Warsaw [1]. Authors elaborate data as a collection of comprehensive results enabling the improvement of science teaching and learning, supporting the creation of legal regulations in the area of education, improving the formula of external examinations as well as teacher education. The discussed data are presented with a special focus on results connected to physics education.

1. Introduction
The reported study was carried out at the Educational Research Institute as part of the project activities called Supporting the realisation of the first stage of the implementation of the Integrated Qualification Systems at the level of central administration and institutions granting qualifications and ensuring the quality of awarding qualifications co-financed from the European Social Fund (ESF) under the Operational Program Knowledge, Education, Development (POWER), Priority II: Effective public policies for the labour market, economy and education, Action 2.13 Transparent and coherent National Qualifications System. The study was conducted in the school year 2017/2018 on one of the last groups of junior high school graduates (16-year-olds), in the year beginning a major systemic change in Polish education – the extinction of junior high schools (gymnasiums) and the associated extension of primary school education by two years. Therefore the findings of this study could serve as a basis for monitoring changes in the level of skills and knowledge of graduates of an 8-grades elementary school, while it should be noted that graduates of the reformed primary school are a population of 15-year-olds. By using only slightly supplemented research tools and the procedures the same as in the first cycle of the Thinking Laboratory in the year 2011 the researchers obtained data allowing them to make a diagnosis of junior high school graduates' skills in science subjects. The report compares the results over the years, but it is not the main purpose of the study and should be treated as a starting point for looking at the findings in a broader cognitive perspective.
2. The main objectives of the study
The aim of the study is to verify students' skills in the subject of natural sciences: biology, chemistry, physics and geography after obtaining qualifications at the first level of the Polish Qualifications Framework and to identify strengths and weaknesses of various aspects of teaching and learning. As the Polish Qualifications Framework is in line with the European Qualifications Framework and frameworks of other countries [2] such an initiative can possibly support also European educational researchers. The survey, carried out on a nationwide sample, will possibly allow comparison of learning outcomes achieved (in terms of knowledge, skills and social competences) with those assumed in documents regulating the work of Polish schools. The primary goal of the study is to diagnose the results of teaching science at junior high schools (gymnasium) and to observe the long-term effects of changes in the education system. The research is aimed at allowing the formulation of recommendations for future work on improving the education system in the area of natural sciences. An important goal of this initiative is also to shape the social reception of educational research as a source of reliable and valuable information on the components of the education system.

2.1. Physics perspective
The authors are particularly interested in findings strongly linked to physics education and having possible impact on teaching and learning processes within this very subject. Therefore the majority of presented findings are those obtained by analysing students achievement in solving physics related problems [3].

3. Characteristics of the research
Field research – Laboratory of Thinking – was carried out in Polish upper secondary education schools in the period from October 2017 to January 2018. The procedures were strictly standardized in order to ensure the comparability of the results of subsequent stages and the smooth conduct of the study. All details of the procedures are described in the contractor's report [4] and only those necessary to understand the presented findings are elaborated in the paper.

3.1. Concept of the research
The research concept is based on the following assumptions:

- cyclicity (after 3 years of the learning cycle) – the period corresponds to full Polish junior high school education.
- random selection of a large number, stratified population of diagnosed students.
- using a significant number of different tasks of certain types.

3.2. Test sample
7288 students from 184 institutions from all over Poland took part in the Laboratory of Thinking research. Upper secondary schools were drawn to participate in the study reflecting the structure of the types of schools to which gymnasium graduates in Poland go. The choice of upper secondary schools, rather than gymnasium, was dictated by the reasons identified in previous studies, and are:

- upper secondary school classes consists of students from various gymnasiums, which significantly reduced the influence of one teacher and thus increased the representativeness of the sample.
- carrying out the survey at the end of the school year in a gymnasium would be burdensome due to the daily duties of employees of these institutions and the fact that students leaving school could be potentially insufficiently motivated to solve test tasks.
- thanks to such a profiled sample, it was also possible to examine the relationship between the level of students' skills and their preferences in choosing the further path of education.
In order to increase the representativeness of the sample a systematic draw with a layering [5] was carried out taking into account the selection of the sample due to specific features:

- type of school (general high school, technical high school, basic vocational school).
- school management body (public school, associated school).
- location (village / city) and size of the respondent's place of residence (<3 thousands, 3-30 thousands, 20-50 thousands, 50-100 thousands, 100-200 thousands, 200-500 thousands, 500-1000 thousands, > 1000 thousands).
- type of school due to gender proportions (balanced, at least 80% boys, at least 80% girls).
- size of the school.

Moreover the draw took into account the proportions of the number of branches in the first classes. The Table 1 illustrates a number of students participating in the study by type of school and gender.

Table 1. Number of students participating in the study by type of school and gender (B – boys, G – girls and none – means that the gender of the respondent could not be identified).

| School type     | High school (lyceum) | Technical college | Basic vocational school | Altogether |
|-----------------|----------------------|-------------------|-------------------------|------------|
| no. of schools  | 82                   | 60                | 42                      | 184        |
| sex             | B        G        none | B        G        none | B        G        none | B        G        none |
| population      | 1255     2092     66     | 1445     1023     46     | 871     458     32     | 3571     3573     144    |
| sum             | 3413     2514     1361   |                     |                       | 7288      |

3.3. Research tools
The main research tool consisted of:

- cognitive test (worksheets) in which multiple choice tasks were included.
- student's questionnaire.

The test worksheets used in the study were designed to evaluate the specific key competences [6] in natural science (biology, chemistry, physics and geography) after the third stage of education in Polish system (gymnasium) but with emphasis on skills directly related to scientific reasoning and scientific research skills [7, 8].

3.3.1. Test worksheets. The test worksheet used in the study were prepared to check the core curriculum skills in science subjects (biology, chemistry, physics and geography) after the third stage of Polish education system [10] but with an emphasis on skills directly related to reasoning and scientific research skills which were given particular attention in the 2009 Polish core curriculum. During the preparation and formulation of subject related tasks a lot of attention was devoted to observation and experimentation. These skills are verified in tests using tasks similar to those appearing in the lower secondary national school examination procedures [11], specifically:

- multiple-choice tasks with one correct answer – the tests consist of tasks containing from 4 to 6 answers to choose from.
"true-false" tasks asking to assign a specific logical value to each statement – the following options are available: yes or no if the student is supposed to answer the question, true or false when the task is to assess the truthfulness of the statements.

"assignment" tasks which include problems with a complex structure, sometimes containing elements of two previously described forms of tasks, but demanding complex actions, e.g. in the first part requiring an answer to a question and in the second an indication of the justification for the choice.

The study used test worksheets in 16 versions, differing in the number of pages, but containing the same number of tasks. The time to solve all problems in the worksheets was no more than 70 minutes. All tests were created as part of the survey conducted in 2011 [12]. In total 208 closed tasks were used in the study, requiring the indication of the correct answer (marking the "x"). The full collection contained 52 tasks in biology, chemistry, physics and geography. Tasks in each subject were divided into four clusters consisting of 13 tasks, selected so that the clusters were comparable in terms of difficulty and time needed to solve them. Subject clusters were then combined in various combinations in pairs to form 16 test worksheets. Each test contained two clusters from two science subjects. A student participating in the study solved tasks from one worksheet, therefore related only to two science subjects. When assigning worksheets to individual students, a distribution model was used to ensure that each of the 16 test variants was completed in a similar number.

3.3.2. Physics example task. As there are further plans to use the research tools in only slightly modified form in future, the students’ tasks could not be revealed. Therefore to illustrate the type of physics problems tackled in the study we use the task presented in the Figure 1 coming from similar work using the same methodology and sharing the same scientific approach [13].

![Figure 1](image_url)

Chose one answer from a.-c. in case I and II properly describing the situation.

I. Tower A exerts on the ground
   a. less pressure than tower B.
   b. the same pressure as tower B.
   c. more pressure than tower B.

II. In such a situation one can use
   a. only equation 1.
   b. only equation 2.
   c. both equations.

3.3.3 Student questionnaire. The student's questionnaire was designed as a separate set of 28 questions of which 23 were closed questions and the others required extended written answers. The student's questionnaire was developed during the first run of the study and underwent slight modifications during their implementation in subsequent runs. The questionnaire questions were selected with a view to expanding the research area and were modeled on the questions of the PISA 2015 survey questionnaire [8, 9]. The data obtained with use of the questionnaire could be categorized anonymously into the following groups:
- characteristics of the respondents themselves (e.g. gender, date of birth, place of residence).
- information about the legal guardians of the respondents (e.g. professional situation, education).
- respondents’ opinions on the gymnasium they attended (e.g. changes in science teachers, atmosphere).
- respondents’ self-perception (e.g. independent thinking, coping with difficult situations, hierarchy of importance of science subjects; valuing sources of knowledge).
- description of educational situations during subject lessons in which respondents participated (e.g. requirements for knowledge of facts, contextual teaching).
- respondents’ perception of elements of science teachers’ professional performance (e.g. occurrence of cross-subject correlations, use of multimedia aids).
- acceptance of respondents for pseudoscientific and other statements.

3.4. Scaling and statistical analysis

In the presented study the Pencil and Paper Interview (PAPI) research technique was used. A detailed analysis of the results was preceded by procedures aimed at determining the level of difficulty of the tasks in the new study group and comparing them with key characteristics of previous studies. Meeting certain requirements resulting from the procedures used was necessary to be able to compare the results obtained and formulate conclusions and recommendations based on them. For this purpose, students’ tasks and results were scaled using the Item Response Theory (IRT) method [14]. The average value was 500, with a standard deviation of 100. Scaling was performed for each science subject separately. As a result of scaling, a six-level skill scale was created in which the student’s achievements were assigned to six levels (I-VI). The ranges of scores given in square brackets shown in the Table 2 were arbitrarily assigned to certain collection of skill identified in the study.

| Table 2. Skill levels with assigned range of scores. |
|---------------------------------|---|---|---|---|---|---|
| Level | I   | II  | III | IV  | V   | VI  |
| score | [360-420] | [420-480] | [480-540] | [540-600] | [600-660] | [>660] |

Statistical analysis of quantitative variables (i.e. expressed in numbers) was performed by calculating the mean, standard deviation, median, quartiles, minimum and maximum. Statistical analysis of qualitative (i.e. not expressed in numbers) variables, both ordinal (those for which a logical order of accepted values can be created) and nominal (those for which no order can be created, e.g. gender), was carried out by calculating the number and percentage of occurrences out of value. Comparison of the values of qualitative variables in the groups was made using the chi-square test (with Yates correction [5] for 2x2 tables) or Fisher’s exact test [3] where the low expected numbers appeared in the tables. The normality of variable distribution was tested using the Shapiro–Wilk test [5]. The comparison of quantitative variable values in two groups was performed using the Student’s t test [15] (when the variable had normal distribution in these groups) or the Mann–Whitney test [15] (when the Shapiro–Wilk test result indicated no grounds for considering the distribution to be normal). Correlations between quantitative and ordinal variables were analysed using the Spearman correlation coefficient [16]. The relationship strength was interpreted according to the following scheme [17]:

- $| r | \geq 0.9$ - very strong relationship.
- $0.7 \leq | r | <0.9$ - strong relationship.
- $0.5 \leq | r | <0.7$ - medium strong relationship.
- $0.3 \leq | r | <0.5$ - weak relationship.
- $| r | <0.3$ - very weak (negligible) relationship.
4. Findings
Data analysis is divided into two parts:

- general analysis of test results in relation to demographic data and level of students’ skills, determined on the basis of the results individually achieved in the study.
- analysis of test results in relation to answers given by student to questionnaire questions.

4.1. General analysis
The Table 3 shows average students’ results obtained in physics in all runs of the Laboratory of Thinking study (standard errors in brackets).

| Year | 2011 | 2012 | 2013 | 2014 | 2017 |
|------|------|------|------|------|------|
| physics | 500,0 (3,9) | 510,9 (4,8) | 499,5 (4,0) | 513,1 (4,0) | 499,9 (3,2) |

There are clear fluctuations in the average results obtained by students in individual subjects without any particular downward or upward trends over the years. The smallest fluctuations were observed in the physics results.

Comparison of test results in physics obtained by boys and girls in the study is presented in Table 4.

| Result | boys | girls |
|--------|------|------|
| mean ± SD | 499 ± 87 | 502 ± 99 |
| median | 500 | 499 |
| quartiles | 435 – 560 | 427 – 577 |

Due to the lack of confirmation of normal distributions in groups the results of boys and girls were compared using the Mann-Whitney test. The groups do not differ significantly in physics results.

The number of students in physics assigned to individual scoring levels is shown in Table 5.

| Level | <I | I | II | III | IV | V | VI |
|-------|----|---|----|-----|----|---|----|
| average | 248 (6,8%) | 548 (15,1%) | 749 (20,7%) | 834 (23,0%) | 684 (18,9%) | 401 (11,1%) | 158 (4,4%) |
| boys | 147 (8,3%) | 259 (14,6%) | 357 (20,1%) | 365 (20,6%) | 328 (18,5%) | 227 (12,8%) | 91 (5,1%) |
| girls | 95 (5,3%) | 273 (15,3%) | 379 (21,3%) | 452 (25,4%) | 349 (19,6%) | 170 (9,5%) | 62 (3,5%) |

Within the physics subject, level III is most frequently represented, while the level below I and level VI are the least numerous. The characteristics of students’ skills associated with those levels are:

- I – student: analyses simple texts; answers direct questions about basic problems using the school’s answer scheme; reads simple information presented in tabular or graphic form; performs simple reasoning with reference to colloquial knowledge (multiple-choice tasks with one correct answer).
III – student: demonstrates basic knowledge of the subject and basic knowledge of the scientific method: distinguishes hypotheses, conclusions and observations, plans a simple experiment and analyses its course, assesses the correctness of simple inference and understands the importance of a control sample; analyses longer texts, including those containing information in graphic form.

VI – student: demonstrates a mastery of subject knowledge, sometimes going beyond the core curriculum; efficiently uses various sources of information, including those given in graphic form (complex thematic maps, charts, diagrams); is fluent in the use of mathematical apparatus for solving tasks; seeking the explanation of natural processes, considering alternative scenarios and adapting research methods to them; when formulating conclusions refers - if necessary - to the category of probability; justifies statements.

Average results obtained by students in physics by school type is presented in Table 6.

**Table 6.** Average results by school type.

| Result                | High school (lyceum) | Technical high school | Basic vocational school |
|-----------------------|----------------------|-----------------------|-------------------------|
| **mean ± SD**         | 545 ± 84             | 484 ± 80              | 416 ± 69                |
| **median**            | 547                  | 482                   | 408                     |
| **quartiles**         | 490 – 605            | 426 – 539             | 368 – 451               |

The analysis showed that high school students achieved significantly higher results in physics than technical high school students which in turn obtained significantly higher results than vocational school students. Thus the results for the subject significantly depend on the type of school. The same trend was observed for all other science subjects. Based on a detailed analysis of the results of each subject it can be stated that the results depend significantly on the place of residence. In physics:

- students from towns with 200-500 thousand inhabitants obtained significantly higher results than students from cities with a population up to 200.000 or over a million.
- students from towns of 500-1000 thousand inhabitants obtained significantly higher results than students from cities of up to 50.000, 100.000-200.000 or over a million.
- students from towns with 50-100 thousand inhabitants obtained significantly higher results than students from towns with up to 20 thousand residents.
- students from towns of 20-50 and 100-200 thousand inhabitants obtained significantly higher results than students from towns of up to 3.000 residents.

Possible causes for the differences might arise from the fact that some students migrate inside and between cities to find a school that fulfils their expectations or they identified a place of temporary residence while completing the questionnaire.

The Table 7 shows comparison of the results for physics tasks used in the study in a light of the effectiveness of their performance by girls or boys. Tasks are divided into six levels of difficulty. The number of physics tasks within each level is indicated by N. There are no tasks at the highest level VI.

**Table 7. Effectiveness of tasks performance by girls or boys.**

| Level                  | <I (N=1) | I (N=3) | II (N=20) | III (N=17) | IV (N=8) | V (N=3) |
|------------------------|----------|---------|-----------|------------|----------|--------|
| Better results – girls | 0 (0,0%) | 0 (0,0%)| 2 (10,0%) | 3 (17,6%)  | 0 (0,0%) | 1 (33,3%)|
| Better results – boys  | 0 (0,0%) | 0 (0,0%)| 1 (5,0%)  | 5 (29,4%)  | 1 (12,5%)| 2 (66,7%)|
| No differences         | 1 (100%) | 3 (100%)| 17 (85,0%)| 9 (53,0%)  | 7 (87,5%)| 0 (0,0%)|
There were no statistically significant differences between genders in the performance on tasks from the two groups with the lowest difficulty level, but there is a clear tendency for boys to achieve some advantage in solving problems with increasing difficulty from level III upwards.

4.2. Analysis in relation to questionnaire answers

In order to analyse the impact of various factors on the results obtained by students the correlation of test results with answers to groups of questions from the student's questionnaire was examined.

One of the questions checked whether the teacher teaching the particular science subjects was replaced by another one during students’ stay at the school (dichotomous scale of answers). The results of such an investigation are shown in Table 8.

**Table 8. Average results with indication of a teacher change.**

| Result  | Yes (N=806) | No (N=2673) |
|---------|-------------|-------------|
| mean ± SD | 509 ± 91   | 500 ± 93   |
| median  | 512         | 499         |
| quartiles | 445 - 574  | 431 - 567  |

The results indicate a statistically significant positive impact of a change of a physics teacher during junior high school education on the results obtained in the study.

Analysis of correlation in the perception of the importance of various school subjects leads to some interesting insights. So called “heat map charts” were used to visualize such correlations. On colorful iconographies in the accepted and consistent convention blue areas indicate strong positive correlations between variables (i.e. the greater the value of one of them, the greater the value of the other). The red areas show strong negative correlations (i.e. the higher the value of one of the variables the lower the value of the other). White areas indicate no correlation. The higher is the color intensity the stronger relationship it illustrates. A heat map showing the correlations between the perception of the importance of coping with six selected school subjects for girls and boys is presented in figure 2.

**Figure 2.** The correlations between the perception of the importance of coping with six selected school subjects for girls (on the left) and boys (on the right) where biologia means biology, chemia – chemistry, fizyka – physics, geografia – geography, język polski – Polish language and matematyka – mathematics.
The emergence of anti-correlation in the girls' perception regarding the importance of coping simultaneously in biology and geography, biology and chemistry as well as chemistry and Polish language was clearly visible but was not observed for boys. For girls, unlike for boys, also a complete lack of correlation between the importance of coping in the study of physics and Polish language was noticed.
Figure 3. Heat maps showing correlations between the perception of the importance of coping with six selected school subjects broken down into different skill levels of students (indicated above each chart).

In the further analysis a correlation between the answers to the question concerning the perception of the importance of coping with certain subject but in the area of different scoring levels achieved by the respondents was examined. As each student solved the tests in two subjects, her/his score was averaged and rounded up if necessary. The findings for all six levels are pictures in the figure 3.

It is clearly visible that as the level of an achievement in the study increases from top to bottom the overall conclusion that it is the "school" as such and not individual subjects that are important to students is weakening. Thus with the level of students’ performance their attitude towards specialization increases. Three clusters of specialization can be distinguished: biology-chemistry (more likely future medical professions), Polish language-geography (future legal professions) and mathematics and physics (future technical, engineering and professional careers in sciences).

These clusters correspond to the specializations indicated in PISA 2015 research in the case of determining what profession students will go into in the future – at the age of 30. In this study the percentage of students pointing to science-related professions totaled over 22%, to engineering and to science-related professions 6%, to medical professions 12% [8, 9].

5. Conclusions

Laboratory of Thinking – diagnosis of teaching science subjects in Poland was an important study picturing teaching and learning science subject in a broad perspective across a number of years. Only some chosen aspects – in authors opinion valuable from physics education point of view – were presented in the paper.

The study shows that the average level of knowledge and skills is not changing significantly across years and is equal among all four science disciplines – the differences between the results from biology, chemistry, physics and geography are statistically insignificant, however it is worth noting that the group of students with the lowest level of skills is dominated by boys. The results depends both on the location of a school and its type.

Unexpectedly the replacement of a physics teacher during education has a positive influence on the students’ performance. This could be associated with a change in the learners’ motivation observed when changes in learning environment take place but this aspect needs further investigation.
In biology and chemistry statistically higher average results are obtained by girls and they score higher than boys in solving most tasks belonging to almost all six levels of difficulty. However, although there is no statistically significant difference between genders in overall results in physics and geography, boys get higher results than girls in difficult tasks in these subjects. Some of the results of the study indicate that the reasons for the above division could be found in standard methods of teaching used in individual subjects preferring more effective teaching of physics or geography in the group of boys and more effective teaching of biology and chemistry in the group of girls. The statement is also supported by results obtained from the analyses of the perception of the importance of certain subjects. As the skill level increases a clear subject specialization is observed.

Perception of the importance of the main subjects of general education – mathematics and mother tongue – strongly depends on the level of skills in the field of physics. Along with the increase of the level of achievement in the test the students’ attitude towards specialization (“deeper but less broader”) increases. Importance of Polish language results goes hand in hand with the validity of results from geography while the results obtained in mathematics remain valid only for people achieving high results in physics.

Acknowledgements
The authors express their gratitude to the Educational Research Institute for enabling the presentation of data obtained as a part of the implementation of the project named Supporting the application of the first stage of implementation of Integrated Qualification System at the level of central administration and institutions that give qualifications and quality assurance which was co-financed by the European Social Fund (ESF) under the Operational Program Knowledge, Education, Development (POWER), Priority II: Effective public policies for the labour market, economy and education, Action 2.13 A transparent and coherent National Qualification System.

6. References
[1] Ostrowska E B, Spalik K (eds.) 2012 Raport z badania. Laboratorium Myślenia. Diagnoza umiejętności gimnazjalistów w zakresie przedmiotów przyrodniczych 2011 (Warsaw: Educational Research Institute)
[2] Find and Compare Qualifications Frameworks. European Qualifications Framework. European Commission. Retrieved 8 November 2016
[3] Good M, Maries A, Singh Ch 2019 Phys. Rev. Phys. Educ. Res. 15, 020129
[4] 2018 Raport realizacyjny z badania Laboratorium Myślenia – diagnoza nauczania przedmiotów przyrodniczych w Polsce (Warsaw/Sopot: PBS Sp. z o.o.)
[5] Agresti A, 2002 Categorical Data Analysis (Hoboken, New Jersey: John Wiley & Sons, Inc.)
[6] Dębowska E, Grechylo T (eds.) 2017 Key Competences in Physics Teaching and Learning Selected Contributions from the International Conference GI REP EPEC 2015 (Springer International Publishing Switzerland)
[7] 2014 PISA 2012 Results in Focus: What 15-year-olds know and what they can do with what they know (OECD).
[8] Federowicz M, Sitek M (eds.) 2017 Program Międzynarodowej Ocen Umiętności Uczniów. Wyniki badania PISA 2015 w Polsce (Warsaw: Educational Research Institute)
[9] Excellence and Equity in Education, summarises student performance in PISA 2015, and examines inclusiveness and fairness in participating education systems, http://dx.doi.org/10.1787/9789264266490-en.
[10] Regulation of the Polish Minister of National Education from 23 December 2008 on The core curriculum for pre-school education and general education in specific types of schools, Dz.U. 2009 No. 4 item 17
[11] Dolata R, Sitek M (eds.) 2015 Raport o stanie edukacji 2014. Egzaminy zewnętrzne w polityce i praktyce edukacyjnej (Warsaw: Educational Research Institute)
[12] Chrzanowski M M, Grajkowski W, Lilpop J, Musialik M, Ostrowska B 2013 Dociekanie naukowe w zadaniach testowych, Nauczanie Przedmiotów Przyrodniczych Kształtujące Postawy i Umiętności Badawcze Uczniów, vol. II (Krakow: Faculty of Chemistry. Jagiellonian University)

[13] Ostrowska E B, Spalik K 2013 Uczymy myślenia. Zadania na lekcje z przedmiotów przyrodniczych (Warsaw: Educational Research Institute)

[14] Reckase M D 2009 Multidimensional Item Response Theory (New York: Springer)

[15] Boslaugh S 2012 Statistics in a Nutshell (O'Reilly Media, Inc.)

[16] Sprent P, Smeeton N C 2001 Applied Nonparametric, Statistical methods, (London: Chapman & Hall/CRC)

[17] Hinkle D E, Wiersma W, Jurs S G 2003 Applied Statistics for the Behavioral Sciences. 5th ed. (Boston: Houghton Mifflin)