Magdaléna Nedelová,
Denisa Šukolová
Slovakia

The New Educational Review

Critical Thinking in Initial Teacher Education: Secondary Data Analysis from Ahelo GS Feasibility Study in Slovakia

DOI: 10.15804/tner.2017.49.3.01

Abstract

The study presents a secondary analysis of the "Assessment of Higher Education Learning Outcomes, Generic Skills Strand", the international study of critical thinking skills (OECD, 2012). National secondary analyses include a comparison of student teachers (n = 110) and students of other study programmes (n = 413) in critical thinking and investigates a potential relation of contextual characteristics with their critical thinking performance. Performance Task and Multiple Choice Questions were used to measure critical thinking skills and a Contextual Questionnaire was used to collect socio-demographic data, and subjectively evaluated characteristics related to test-taking motivation and coursework. The main results showed that student teachers scored lower in critical thinking performance than students of other study programmes.

Keywords: *critical thinking, initial teacher education, assessment, development*

Introduction

In recent years, increased attention has been paid to general intellectual or cognitive competences and skills, which, even if they are not directly tied to a particular curriculum or course of study, are thought to be the salient outcome of postsecondary education. The emphasis on critical thinking (CT) in today's global educational policy results from rapid social changes in recent decades. Expansion of scientific knowledge along with informational and technological

revolution substantially amplified the volume and accessibility of information, as well as demands for their cognitive processing. This requires adopting new strategies of searching for and selecting information, critical evaluation of the content and verifying the reliability of sources. Recognition of the essential role of CT predicts significant changes also in teaching and learning.

Definition of Critical Thinking

Critical Thinking (CT) skills are often referred to as higher order cognitive skills which are relatively complex; they require judgment, analysis, and synthesis, and are not applied in a rote or mechanical manner (Halpern, 1998). Definitional boundaries for CT are fuzzy, but there is a wide expert consensus regarding its defining central features. Most attempts to define and measure CT operationally focus on an individual's capability to do some or all of the following: identify central issues and assumptions in an argument, recognize important relationships, make correct inferences from data, deduce conclusions from information or data provided, interpret whether conclusions are warranted based on given data, evaluate evidence or authority, make self-corrections, and solve problems (Erwin, 2000). In 1990, a cross-disciplinary international panel of 46 experts completed two-year, multi-round, strict-method Delphi research, which resulted in a robust conceptualization of CT for the purposes of instruction and educational assessment. CT is defined as 'purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based' (Facione, 1990). Researchers have distinguished between CT skills and dispositions to CT, suggesting a meaningful distinction between the ability to think critically and the willingness to actually do so (Bailin, 1999; Facione, 2000; Ennis 1996; Halpern, 1999). According to Facione (1990), the core set of CT skills comprises interpretation, analysis, evaluation, inference, explanation and self-regulation. Dispositions, which can be seen as attitudes or habits of mind, include open- and fair-mindedness, inquisitiveness, flexibility, inclination to seek a reason, desire to be well-informed, and respect for and willingness to entertain diverse viewpoints (Lai, 2011).

State of Critical Thinking in Higher Education

Many researchers working in the area of CT call attention to the poor state of CT among tertiary students (Halpern, 1998). Although traditional and contemporary theories provide a base for teaching for CT, many schools are still producing students who are not sufficiently equipped with higher order cognitive skills.

One of the assumptions found in the literature is that CT, although regarded as essential, is often misunderstood (Thompson, 2011).

According to Paul (1992), most universities do not act within the framework of a substantial concept of CT. However, they mistakenly believe they are doing so and assume that students are taught CT. In reality, the transmissive instructional approach prevails and largely ineffective short-term study habits are still the standard in college instruction and learning. In her review, LaPoint-O'Brien (2013) also points out that effectiveness of teachers, or a lack thereof, probably has to do with the high variability of how the CT concept is understood among them. For instance, an older study by Paul, Elder and Bartell (1997) revealed that instructors lack proper CT understanding and practice, and cannot teach it properly. Most teachers declared CT development to be one of their main teaching goals. However, observation showed that only 10% of those teachers were actually able to do it efficiently. A similar study was also carried out by Thomas (1999), who found out that a lot of teachers lacked the necessary vocabulary to discuss CT standards in students. Zohar and Dori (2003) pointed out that many teachers believe that CT can only be developed by high-achieving learners; prejudice towards low-achieving learners resulted in personal demotivation and frustration in students.

If it is broadly believed that CT skills are teachable and learnable, then those who teach others to think critically have to be accomplished critical thinkers themselves. Moreover, in order to make reasonable decisions in relation to curriculum and pedagogical practices, teachers themselves must be able to analyse, criticize and advocate ideas, and to reach judgmental conclusions based on sound inferences. Based on the above, we believe it is particularly important to investigate the level of CT among Slovak teacher students. Accordingly, the aim of this study was (1) to compare the CT level achieved by students of teaching and non-teaching study programmes; and (2) to identify contextual characteristics related to the performance in CT. Secondary analysis of Slovak students' results in AHELO GS might reveal valuable information on the current state of CT in student teachers and outline the way to improve it.

Methods

In AHELO GS feasibility study, two different assessment instruments were used to measure CT skills: Performance Task (PT) and Multiple Choice Test (MCQ). In Slovakia, AHELO GS testing took place in the course of April and May 2012. It was administered through an internet-based test platform. Participants entered

the exam through a secure browser that locked down computer functions and distributed a 90-minute PT and a 30-minute MCQ section to each student. In the last step, each participant completed a 10-minute Contextual Questionnaire (CQ).

Performance tasks (PT) required students to answer several open-ended questions about hypothetical, yet realistic situations. Each PT contained its own document library that included a range of information sources (e.g., letters, memos, research reports, articles). To answer the questions successfully, students were expected to gain a deep insight into the text content, understand the relationships among information items, compare and integrate information from different parts of the text, evaluate the credibility of the source, draw a conclusion on the probable cause of the problem, propose a solution and craft an argument with relevant and reliable information. One of two variants of PT was randomly distributed to each student. The first PT variant (Lake to River) cannot be published due to the copyrights. The second one (Catfish) is publicly available in the AHELO Feasibility Study Report (OECD, 2012, pp. 220–233). PT does not have clear-cut right or wrong answers. Since every prompt can include various possible arguments or relevant information, scorers received prompt-specific guidance in addition to the scoring rubric, which allows for criterion-referenced interpretations of the scores (cf., performance levels description OECD, 2012, pp. 234–236). All responses were double-scored by specially trained scorers on a 6-point scale (1 – lowest quality, 6 – highest quality) within the three dimensions – AR – analytical reasoning and evaluation, PS – problem solving, WE – writing effectiveness. In our research, the total score in PT represents a sum of average scores assigned by two scorers in the three assessed dimensions (scale 0–18).

Multiple Choice Questions (MCQ) required students to select the right answer out of four possibilities. The item stem consisted of several paragraphs of text combined with tables or charts. One of the four MCQ variants was randomly assigned to each student. The total score in MCQ corresponds to the number of questions correctly answered. Since the MCQ variants were not equivalent either in the number of items, or in their difficulty (cf., Nedelová, 2015), only scores in core items (common in all the four variants) were taken into account. To allow for a comparison with the total score in PT, the original MCQ core scale 0–13 was transformed into 0–18 points in our analysis.

Contextual Questionnaire (CQ) collected data about individual students and their experience in the course of their bachelor degree study. Our analysis focused only on selected characteristics related to test-taking motivation and coursework (subjectively assessed on a 4-point scale; 1 – minimum, 4 – maximum).

Research Sample

Sixteen Slovak higher education institutions participated in the AHELO GS feasibility study. The research sample consisted of 1,544 students in the final year of their bachelor degree studies (60% of them female, 75% of them between 21 and 23 years of age). Only volunteers took part in the testing, and the overall participation rate was 58%. In the context of our analysis, we included only the universities which provide (among others) teaching study programmes. Students with blank or entirely off-topic responses to PT, as well as those who spent less than 20 minutes on PT, were not included in our results. These responses indicated low motivation for testing rather than an actual level of CT. A secondary analysis was conducted on the sample of 523 students (110 teaching students, 413 students of other study programmes).

Results

In the first step, we analyzed differences in CT measures between the student teachers and the students of other study programmes. In both groups of participants the observed variables met the requirement of normal data distribution (Table 1). To evaluate the differences, the parametric T-test for two independent selections was used. The results are listed in Table 2. Based on statistical analysis, it can be stated there was a significant difference between the student teachers and the students of other study programmes in PT (Performance Task) – Writing Effectiveness (WE) (p=0.04, t=2.063), and in the PT summary score (p=0.039, t=2.078). The students of other study programmes scored higher than the teacher students; (AM=2.79, AM=9.00 and AM=2.61, AM=8.46). Another significant difference (p=0.010, t=2.574) was observed in the MCQ scores. The students of other study programmes (AM=7.95) scored higher than the student teachers (AM=7.04). For Analytic Reasoning (AR) and Problem Solving (PS) sub-scores, no statistically significant differences between the groups were observed.

Table 1. Descriptive statistics of CT measures in student teachers and students of other study programmes

n=110	PT_summ	PT_AR	PT_PS	PT_WE	MCQ
AM	8.46	2.70	2.70	2.61	7.04
MDN	9.00	3.00	3.00	2.50	6.92
SD	2.41	0.90	0.88	0.81	3.07
Skewness	0.28	0.14	0.29	0.19	0.57

n=110	PT_summ	PT_AR	PT_PS	PT_WE	MCQ
Kurtosis	0.09	-0.13	-0.23	-0.11	0.51
n=413	**PT_summ**	**PT_AR**	**PT_PS**	**PT_WE**	**MCQ**
AM	9.00	2.84	2.84	2.79	7.95
MDN	9.00	3.00	3.00	3.00	6.92
SD	2.44	0.82	0.85	0.84	3.35
Skewness	0.03	-0.07	0.07	0.11	0.47
Kurtosis	-0.26	-0.26	-0.36	-0.33	-0.41

PT_summ – overall score in Performance Task; PT_AR – score in Analytic Reasoning, PT_PS – score in Problem Solving; PT_WE – score in Writing Effectiveness; MCQ – score in Multiple Choice Questions (core items)

Table 2. Values for differences in CT measures between student teachers and students of other study programmes

	PT_summ	PT_AR	PT_PS	PT_WE	MCQ
Student teachers (n=110)	t=2.078*	t=1.607	t=1.589	t=2.063*	t=2.574**
Students of other study programmes (n=413)	p=0.039	p=0.109	p=0.113	p=0.040	p=0.010

PT_summ – overall score in Performance Task; PT_AR – score in Analytic Reasoning, PT_PS – score in Problem Solving; PT_WE – score in Writing Effectiveness; MCQ – score in Multiple Choice Questions (core items), *- statistical significance of $p \leq 0.05$; ** – statistical significance of $p \leq 0.01$; t – resulting value of T-test

In the second step, we analyzed contextual characteristics related to CT measures. In the description of the variables, we focused particularly on the indicators of data distribution. Since in a number of cases indicators of skewness and kurtosis exceeded the -1 to 1 interval, the Spearman Correlation Coefficient was used to reveal the correlations between them. Correlations were examined on the whole set of participants (N=523). In Table 3 the overview of statistically significant results can be found.

Based on the analysis, it can be stated that the CT measures (overall score in PT; AR, PS, WE sub-scores; MCQ score) are in a significant positive correlation with: (1) relevance of the test to the students' field of study; (2) relevance of the test to the students' future professional practice; (3) effort put into the test; (4) involvement in classroom discussions; and (5) coursework emphasis on memorising. In all the cases, the correlations are weak – $\rho \leq 0.30$. It can be further stated

that the CT measures are in a significant negative correlation with coursework emphasis on: (1) synthesizing; (2) making judgements; (3) applying; (4) designing new products, and (5) working in groups/teams with other students. In all the cases, the correlations are weak – $\rho<0.30$.

Table 3. Values for correlations between contextual characteristics and CT measures

Spearman's rho	RFS	RFP	EFF	MEM	SYN
PT_summ	ρ=0.102* p=0.020	ρ=0.169*** p=0.000	ρ=0.124** p=0.005	-	-
PT_AR	-	ρ=0.138** p=0.002	ρ=0.122** p=0.005	-	-
PT_PS	-	ρ=0.141*** p=0.001	ρ=0.121** p=0.005	-	-
PT_WE	ρ=0.104* p=0.017	ρ=0.146*** p=0.001	ρ=0.117** p=0.007	-	-
MCQ	ρ=0.157*** p=0.000	ρ=0.159*** p=0.000	-	ρ=0.093* p=0.021	ρ=-0.101* p=0.021
Pearson's rho	**MJ**	**APP**	**DNP**	**WGT**	**ICD**
PT_summ	-	-	-	ρ=-0.136** p=0.002	ρ=0.105* p=0.017
PT_AR	-	-	-	ρ=-0.124** p=0.004	ρ=0.089* p=0.041
PT_PS	-	-	-	ρ=-0.111* p=0.011	ρ=0.118** p=0.007
PT_WE	-	-	ρ=-0.106* p=0.015	ρ=-0.116** p=0.008	ρ=0.107* p=0.014
MCQ_core	ρ=-0.095* p=0.030	ρ=-0.103* p=0.019	ρ=-0.121** p=0.006	ρ=-0.114** p=0.009	-

RFS – relevance to the field of study; RFP – relevance to the future professional practice; EFF – effort put into the testing; MEM – coursework emphasis on memorizing; SYN – coursework emphasis on synthesizing, MJ– coursework emphasis on making judgements, APP – coursework emphasis on applying, DNP – coursework emphasis on designing new products, WGT – coursework emphasis on working in groups/teams, ICD – involvement in classroom discussions * – statistical significance of $p\leq0.05$, ** – statistical significance of $p\leq0.01$

Discussion

In general, 34% of the subjects performed at levels 1 and 2 (very poor or poor answer; cf. Scoring Rubric). The analytical parts of their answers often included only literal rewriting of short paragraphs from the provided texts, unrelated to the actual question. Due to overall disorganization of the text, the argument was untraceable or had a simplistic structure. Written expressions were fragmentary and hard to understand. The largest group of students (45%) reached the performance level 3. Their answers were brief, which can be related to communication standards in the Slovak educational culture (open test questions require brief and clear answers). In the analytic part of the questions, fragments of the text appeared. However, in this case the answers were usually related to the questions. The argument was relevant and logical; however, it was usually limited to a single aspect of the problem, and was not elaborated on. Performance at level 4 (and higher) was reached only by 21% of the Slovak students. Here, independent interpretations and inferences across the whole answers were employed.

The results indicate that despite its importance and broad relevance, CT is a neglected topic in the Slovak education system, which has been documented also in the TIMMS and PISA studies (NÚCEM, 2013; OECD, 2015). According to the study reports, Slovak pupils know how to use procedures accurately, but they lack a deep understanding of the internal meaning of concepts and procedures and, as a result, they often fail to use the concepts and procedures as disciplinary tools to accomplish meaningful goals in the world. In higher education, students very often despise CT exercises or perceive them as difficult, because they are not trained to probe, question, or analyse in the course of their earlier studies.

Critical Thinking Level in Student Teachers and Students of Other Study Programmes

Based on the above-mentioned overview, the CT level of a relatively large portion of Slovak students (regardless of their study programme) is below average. The students of other study programmes scored significantly better than the student teachers in both PT and MCQ. In a closer view, constructing arguments in a coherent way (writing effectiveness, WE) as a part of PT sub-scores proved to be particularly challenging for the student teachers. Since the language is the main tool used by teachers in their profession, this finding is alarming. This can be probably explained by the national and cultural particularities. Since the current Slovak education system can be generally considered as unable to properly foster higher order cognitive skills, we believe that these differences result from individ-

ual dispositions of students. Based on this assumption (not low quality of initial teacher education is a cause of this phenomenon), the problem can be tracked down to teacher salaries in Slovakia. Here we are speaking of a long-term issue; Slovak teachers earn considerably less money than their colleagues from other EU countries (EURYDICE report, 2013/2014). This has a significant impact on the social status of the profession, which subsequently influences the academic level of students entering higher education.

Contextual Characteristics related to Critical Thinking Measures

In the second part of the research, the correlation between the students' performance in CT (PT, MCQ) and selected contextual characteristics were examined. We expected that the identified positive correlations could indicate study components which can be considered as supportive factors of CT. Counter-intuitively, the data show a positive correlation of CT performance with the coursework emphasis on memorizing, and a negative correlation of CT performance with the emphasis on synthesizing, making judgements, applying and designing new products and working in groups/teams. A possible explanation can be that the students with more developed CT are more sensitive to the undersized aspects of their study and excessive memorizing typical in formal education. Using criteria (having higher expectations) different from their counterparts, the students with more developed CT might systematically lay a higher emphasis on memorizing, and lower emphasis on higher order cognitive skills, applying, and collaborative learning activities in the course of their study.

A positive correlation of the CT score with the effort put into the test, relevance of the test to the students' field of study and their future profession is not surprising. The effort put into the test is clearly an indicator of the students' test motivation, which determines their rate of success regarding cognitively demanding tasks. Moreover, the level of the students' test motivation is higher when the test is subjectively perceived as relevant to their field of interest. This also suggests that the context of the particular PT might serve as a confounding variable. For instance, the Catfish PT would be perceived as familiar for students of natural sciences rather than for students of different fields of study such as arts or humanities. Finally, there is a positive correlation of CT performance and the students' involvement in classroom discussions. This result is consistent with the research findings reported in Pascarella's and Terenzini's meta-analysis (2005).

Conclusions

In the secondary analysis of the data collected in the AHELO GS feasibility study, we gained a deeper insight into the quality of the initial teacher education, and broadened the knowledge in this field, which still requires lots of research. To improve students' performance in CT, educational institutions must improve the initial teacher education – integrate CT skills into all aspects of future teachers' training and train them to be models of effective thinking strategies. A clear scope for further research is evident, it is still necessary to examine the extent to which CT is embedded in various subject areas of current teacher education courses (employing analysis of assignment topics, marking guides, subject outlines, learning materials, etc.).

References:

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999).Conceptualizing critical thinking. *Journal of Curriculum Studies,* 31(3), 285–302.

EACA EURYDICE (2015). *Teachers' and School Heads' Salaries and Allowances in Europe, 2013/14*. Retrieved from: http://eacea.ec.europa.eu/education/eurydice/documents/facts_and_figures/salaries.pdf

Ennis, R. (1996). Critical thinking dispositions: Their nature and assessability. *Informal Logic, 18* (2&3), 165–182.

Erwin T. D. (2000). *The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing*. U.S. Department of Education. Washington, DC.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.* CA: The California Academic Press.

Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement and relationship to critical thinking skills. *Informal Logic,* 20 (1), 61–84.

Halpern, D. F. (1998).Teaching critical thinking for transfer across domains: Dispositions, skills, structure training and metacognitive monitoring. *American Psychologist,* 53 (4), 449–455.

Halpern, D. F. (1999). Teaching for critical thinking: Helping students develop the skills and dispositions of a critical thinker. *New Directions for Teaching and Learning,* 80, 69–74.

Lai, E. R. (2011). *Critical Thinking.* Assessment & Information group of Pearson. Retrieved from: http://images.pearsonassessments.com/images/tmrs/CriticalThinkingReviewFINAL.pdf.

LaPoint-O'Brien, T. (2013).Action Research: The Development of Critical Thinking Skills.

ERIC Online Submissions. Retrieved from: http://files.eric.ed.gov/fulltext/ED540359.pdf.

NÚCEM (2013). Národná správa **z** medzinárodných výskumov PIRLS 2011 a TIMMS 2011. Retrieved from: http://www.nucem.sk/documents/20131015_Klucove_kompetencie_web.pdf

OECD (2015). PISA 2015 – country overview. Retrieved from: http://www.compareyourcountry.org/pisa/country/svk

OECD (2012). *AHELO Feasibility Study Report, Volume 1, Design and implementation.* Retrieved from: http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf.

Nedelová, M. (2015): *Assessment and Development of Critical Thinking and Argumentation Skills of Bachelor´s Degree Students.* (PhD thesis). Banska Bystrica: PF UMB, p. 127.

Pascarella, E. T. & Terenzini, P. T. (2005). *How College Affects Students, Volume 2, A Third Decade of Research.* San Francisco, CA: Jossey-Bass, 155–212.

Paul, R. W. (1992). Critical thinking: What, why, and how? New Directions for Community Colleges, 1992 (77), 3–24.

Paul, R., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations.* ERIC Document #437 379.

Thomas, P. (1999). *Critical thinking instruction in selected greater Los Angeles area high schools.* Doctoral dissertation, School of Education and Behavioral Studies, Azusa Pacific University, Azusa, CA.

Thompson, C. (2011). Critical Thinking across the Curriculum: Process over Output. *International Journal of Humanities and Social Sciences.* 1 (9), 1–7.

Zohar, A. & Dori, Y. (2003). Higher-order thinking skills and low-achieving students: Are they mutually exclusive? *The Journal of the Learning Sciences*, 12 (2) 145–181.