BABEŞ-BOLYAI UNIVERSITY

FACULTY OF PSYCHOLOGY AND EDUCATIONAL SCIENCES

DOCTORAL SCHOOL 'EVIDENCE-BASED PSYCHOLOGICAL ASSESSMENT AND INTERVENTIONS'

Ph.D. THESIS SUMMARY

COGNITIVE PREDICTORS OF ACADEMIC EXCELLENCE: CRITICAL THINKING AND METACOGNITION

AUTHOR: Ph.D. CANDIDATE ANDREA BARTA

SCIENTIFIC ADVISOR: PROFESSOR Ph.D. ISTVÁN SZAMOSKÖZI

Cluj-Napoca

2023

ACKNOWLEDGEMENTS

This doctoral thesis is the result of my own effort and the contribution of several valuable people. In the following I would like to express my thanks and gratitude to the people who helped me with the elaboration and finalisation of this thesis from professional and personal points of view. First of all, I would like to express my sincere gratitude to my scientific adviser, Professor István Szamosközi, Ph.D., who introduced me into the research process and raised my attention to it, who supported me, encouraged me from the beginning of our common work, provided me with very useful professional advice, guided me in the research and writing process of the thesis, transmitted me extensive theoretical and methodological knowledge that I utilised in my doctoral studies. I am grateful to him for promoting both my professional and personal development. Furthermore, I would like to express my thanks and gratitude to the members of my supervision committee, namely Professor Aurora Szentágotai-Tătar, Ph.D., Professor Anca Dobrean, Ph.D., and Assistant Professor Silviu Matu, Ph.d., for their substantial support, advice and mentorship. I would also like to express my appreciation to the professors beyond my supervision committee and to my colleagues from the Doctoral School of Evidence-Based Assessment and Psychological Interventions and the Department of Applied Psychology, for their professional and emotional support and for all the stimulating discussions. Finally, I am sincerely thankful for the support and encouragement I received from my family and friends during my doctoral studies.

Notes.

(1) This note is to certify by Andrea Barta that:

(a) The thesis includes the original research work of Andrea Barta towards the Ph.D.

(b) Parts of the thesis have been accepted or submitted for publication or presented as conference papers; appropriate citations for these publications were included in the thesis. Other co-authors have been included in the publications if they contributed to the preparation of the published text, data interpretation, etc. (their contribution was clearly explained in the footnotes of the thesis).

(c) The thesis was written according to academic writing standards. All the text of the thesis and its summary was written by Andrea Barta who assumes all the responsibility for the academic writing; also:

- A software was used to check for academic writing (see at http://www.turnitin.com/); the thesis has passed the critical test;
- A copy of the research dataset/database was delivered to the Department/Graduate School.

Signature for certifying the Notes: Ph.D. Candidate Andrea Barta.

(2) All the Tables and Figures are numbered within the corresponding chapter or subchapter of the thesis.

Table of Contents

Keywords: academic excellence, cognitive processes, metacognition, critical thinking, competition, Olympiad.

CHAPTER I. THEORETICAL BACKGROUND

1.1. Extracurricular Activities and Academic Achievement as the Main Indicators of Academic Excellence

One criterion of academic excellence is high learning performance, but the expertise in certain scientific, sports or art areas also contributes to the development of those 21st century skills, higher-order cognitive abilities, which assure the eminent performance of students. During primary and secondary school years these activities could be competitions, Olympiads, during academic years eminent students could participate in conferences organised by universities. The experience and new skills acquired through extracurricular activities are transferred by students into the learning process, so these activities can also contribute to the increase of their learning achievement. Several studies highlighted the positive association between the attendance in extracurricular activities and learning performance (Feraco et al., 2022; Fujiyama et al., 2021; Kravchenko & Nygård, 2022; Lang, 2021).

The positive effect of extracurricular activities on learning achievement can be explained by three theoretical approaches: the social control theory, the social learning theory and the social capital theory. Social control is present in extracurricular activities, namely the participation in regular activities assumes compliance with the pre-defined rules of the institution, for which students could receive institutional reinforcement, reward. Based on the social learning theory, students could learn from their better achieving peers, from their teachers or from other experts, which also has a positive effect on their learning results. The social capital theory emphasises the role of extracurricular activities in forming relationships not only with student peers, but also with parents and educational staff members, which contributes to a more effective information flow and support from the people with various positions in the institutional hierarchy. As opposed to extracurricular activities, close relationships and friendships have a symmetrical effect on learning performance because these friendships could have positive but also negative effects on achievement. Nevertheless, relationships formed during extracurricular activities have an asymmetrical effect, since extracurricular activities amplify the relationships that have positive effects on the learning achievement and weaken damaging friendships, in the case of low achievement students (Fujiyama et al., 2021).

1.2. Critical Thinking as an Explaining Cognitive Ability of Academic Excellence

1.2.1. Theoretical Models of Critical Thinking

In the literature, there is no coherent definition of critical thinking (Paul, 2004). Critical thinking involves rational, targeted, planned thinking. Definitions generally include scientific deductions, evaluation of statements, recognition of persuasion, generation of options and alternatives. Yanchar et al (2008) identified analytical deduction and discovery as key components of critical thinking. Critical thinking requires deep and meaningful understanding of information, overcoming fallacies and errors (Halpern & Sternberg, 2020).

In the late 1980s, the APA Delphi study aimed to provide a unified definition of critical thinking with the use of the concept analysis method. The Delphi research method aimed to review the basic issues and definitions of critical thinking. More than half of the experts who participated in the study were philosophers, but also teachers, social scientists and physicists shared their professional experience and arguments about critical thinking. The process was led by a director, with participants providing detailed written answers supported by arguments to six sets of questions, which the

director shared with everyone. The experts were allowed to see each other's opinion anonymously, so they did not influence each other's decisions. More than 95% of the experts agreed that critical thinking skills include analysis, evaluation and inference, and a majority also agreed that the core skills of critical thinking include interpretation, explanation and self-regulation. Although critical thinking skills are present in many domains, i.e. they are domain-general skills, their application in different contexts requires domain-specific knowledge, using different methods to make informed decisions. The study put particular emphasis on the identification of critical thinking dispositions alongside critical thinking skills, as the philosophers involved in the research drew attention to the personality traits that determine critical thinking, in the absence of which critical thinking skills are not applied (Facione, 1990).

1.2.2. Critical Thinking: Skills and Attitudes

Critical thinking is a combination of skills, attitudes and knowledge, which includes critical thinking skills and dispositions as well (Ennis, 2018; Facione, 1990; Haber, 2020). In order to achieve it, it is not sufficient to possess high levels of critical thinking skills alone, but it also requires the presence of certain motivational and personality factors that contribute to the application of these skills in practice (Ennis, 2018; Halpern & Sternberg, 2020).

Facione's (1990) theoretical approach emphasises six different comprehensive cognitive skills: interpretation, analysis, evaluation, conclusion, explanation, self-regulation. Interpretation involves categorisation, decoding, significance and clarification of meaning. The second skill is analysis, which involves the examination of ideas, formulating arguments and their analysis. Evaluation as a cognitive skill is nothing more than the assessment of claims and evidence. Conclusion involves questioning evidence, hypothesising alternatives and drawing conclusions. Explanation involves stating results, justifying procedures and presenting evidence. Self-regulation is the sixth skill of critical thinking, which involves self-assessment and self-improvement as well.

Affective dispositions are strategies that include the development of intellectual humility and decision delay, intellectual curiosity, perseverance, and the utilising of confidence in arguments. A disposition for critical thinking implies engaging in complex tasks and persevering in these tasks, flexible thinking, and the ability to abandon unproductive strategies and correct mistakes when necessary (Halpern & Sternberg, 2020). Van Dongen et al. (2005) concluded that a disposition for critical thinking enables a person facing uncertainty and conflict to solve problems and make decisions by asking questions, considering alternatives and conflicts. The critical thinking disposition includes motivation and a person's commitment to apply critical thinking in necessary situations (Halpern, 1997). The affective dimension can be further subdivided into two subgroups, motivational tendencies and behavioural habits. Motivational tendencies are the tendency to know the truth and a disposition of curiosity, while behavioural habits include analytical and compliance habits (Genç, 2017).

1.2.3 Critical Thinking and Academic Performance

Critical thinking is a predictor of academic excellence and school success, as it contributes to a more accurate mastery and transfer of the learning material by analysing the material, identifying claims, interpreting and evaluating information, and drawing deductive and inductive conclusions. Several research studies with secondary school and university students have found a positive relationship between

critical thinking and academic performance (Ali & Awan, 2021; Ibrahim et al. 2021; Ng et al., 2022; Ren et al., 2020; Shahzadi et al., 2020).

Akpur (2020) investigated the extent to which undergraduate students' higher-order cognitive processes such as creative, reflective and critical thinking are interrelated and predictive of academic performance. As a result, the research found that the three cognitive processes are antecedents of each other and all three significantly and positively predict academic performance, measured by the annual overall average of grades.

Ren et al. (2020) investigated the extent to which critical thinking skills and dispositions, along with general cognitive abilities, predict academic performance. They identified fluid intelligence, working memory and processing speed as general cognitive abilities. The relationship between critical thinking, general cognitive abilities and academic performance was also examined in two different samples of primary school and university students. After controlling general cognitive ability, both critical thinking skills and dispositions predicted academic performance. When the predictive power of critical thinking and general cognitive abilities on academic performance was examined in a single model, the effect of critical thinking abilities largely overlapped with general cognitive abilities, whereas the individual predictive power of critical thinking dispositions was significant.

Doleck et al. (2017) examined the relationship between computational thinking skills such as cooperativeness, creativity, algorithmic, critical thinking, problem solving and academic performance in Canadian high school students but found no significant relationship between the measured skills and academic performance, with the exception of cooperativeness.

1.2.4. Critical Thinking and Extracurricular Activities

Participation in extracurricular activities such as competitions, conferences, academic and social group memberships contribute to the development of critical thinking skills and dispositions (Chen et al., 2020; Jatmiko et al., 2020; Schiefer et al., 2020). Othman et al. (2015) investigated the impact of a three-week debate competition on the critical thinking of 16-year-old English as a foreign language learners. They found a significant difference in critical thinking measured in pre- and post-tests. After the debate activity, they scored higher in their critical thinking skills. Critical thinking skills are developed through collaborative learning during the debate, and reasoning, questioning, drawing conclusions and explaining are also critical thinking skills that students actively apply during debate activities.

Some research has highlighted that participation in mathematics competitions and mathematical problem solving are significant predictors of critical thinking components (Utomo, 2018; Wahidin & Romli, 2020). Other research has concluded a positive relationship between participation in chemistry competitions, robotics competitions and students' critical thinking (Chen et al, 2020; Jatmiko et al., 2020; Nugroho et al., 2019). Hong et al. (2013) investigated the developmental effects of an after-school robotics program. The science and technology program had positive effects on student collaboration, learning and creativity. Merino-Armero et al. (2021) investigated the impact of extracurricular robotics classes on the computational skills of Spanish fifth-grade students, controlling their fluid intelligence. Significant differences in computational abilities were found between students who did not participate in after-school robotics classes and those who participated in these programs for 2 or more years.

1.2.5. The Applicability and Development of Critical Thinking in Education

Butler et al. (2017) examined in their research whether intelligence or critical thinking ability play a more important role in everyday decisions. Their results suggest that critical thinking is more closely related to decisions than intelligence, leading them to conclude that critical thinking predicts what people do in real life.

Psychology is an evolving discipline, new theories, practices and methods are emerging, so students, future psychologists, need to be constantly informed, to acquire new information, but being exposed to large amounts of information requires a developed level of critical thinking skills and their application. When learning psychology, students apply their critical thinking skills by identifying which theoretical models support, complement or contradict each other, which theoretical models are supported by scientific arguments, which requires their reasoning and fallacy detection skills. Critical thinking is also necessary to identify which credible electronic sources or journals are worthy of attention. Psychological research requires critical thinking skills such as the ability to test hypotheses within inductive reasoning to provide correct support for hypotheses, and the ability to design experiments to help formulate hypotheses correctly and in detail, design experimental and control groups, and keep randomisation in mind (Ennis et al., 2005). Critical thinking is also important in exploring cause and effect, connecting theory and practice, designing research plans and conducting research (Ho et al., 2014).

The development of critical thinking can be achieved through methods such as the use of concept maps (Carvalho et al., 2020; Mohammadi et al., 2019; Roshangar et al., 2020; Tseng, 2019), a learning and teaching strategy that can be used to develop situational, context-independent, analytical thinking (Ligita et al., 2022) and problem solving (Wang et al., 2018; Zwaal & Otting, 2012), as it promotes deep processing of information, as opposed to superficial processing (Garwood et al., 2018). Concept mapping, similarly to critical thinking, is a nonlinear cognitive activity that is suitable for developing high-level cognitive processes such as critical thinking and decision-making (Alfayoumi, 2018; Khrais & Saleh, 2020).

1.3. Metacognition as an Explaining Cognitive Ability of Academic Excellence

1.3.1. Theoretical Models of Metacognition

Metacognitive processes and strategies are essential in university education, as students have to process large amounts of information from different sources. The study and processing of the learning material requires the use of metacognitive regulatory processes that allow the organisation of the learning process, time management, organisation of information, monitoring of comprehension, identification of errors in one's own performance, their correction, and evaluation of individual abilities and performance (Karatas & Arpaci, 2021).

Metacognition can be defined as thinking about cognitive processes and thinking (Flavell, 1987), and applying cognitive processes and strategies appropriately, in the right place, in the right way, at the right time. We can distinguish between two main components of metacognition, knowledge about cognition and the regulation of cognition, or metacognitive experience (Dindar et al., 2020; Flavell, 1987; Schwarz et al., 2021). Metacognitive knowledge refers to general knowledge about cognitive processes, the recognition of cognitive strengths and limitations, and the assessment of the abilities of others. This knowledge is acquired through experience and is generally stable. Metacognitive knowledge includes knowledge such as the ability to distinguish between an easier and a more difficult text, and the ability to identify effectively which strategies to use to interpret texts of different difficulty. Metacognitive experience involves the evaluation and regulation of cognitive

processes and activities that are not necessarily stable. When students ask themselves questions while reading, they are evaluating their understanding of the text, regulating their cognitive processes. When students decide to reread one or more sentences or a paragraph because they have difficulty understanding a particular passage, they regulate their comprehension. Metacognitive knowledge can lead to increased metacognitive experience and performance, but it does not always do so. Knowing that a complex message should be re-read for better understanding does not guarantee that it will be done so. In addition to metacognitive knowledge, motivation is also required to perform certain tasks (Dindar et al., 2020).

Knowledge about mental processes and operations consists of three subcomponents, which include the reflective aspect of metacognition. Declarative knowledge comprises knowledge about the mind and cognitive strategies, procedural knowledge refers to the application of declarative knowledge in practice, while conditioned knowledge refers to the knowledge about when and how to apply particular strategies (Smith et al., 2020; Vianin, 2016).

The regulation of cognition involves sub-processes that facilitate learning control. Research has identified five main regulatory skills: planning, information management strategies, monitoring comprehension, error detection strategies and evaluation (Mowling & Sims, 2021; Smith et al., 2020). Metacognitive regulatory skills are essential in the lifelong learning process (Bransen et al., 2022).

1.3.2. Metacognition and Self-Regulated Learning

According to Zimmerman (1990), metacognition involves the complex interaction of psychological variables such as motivation, emotions and behaviour. If these factors were ignored, the overall lack of correlation between academic performance and metacognition, metacognitive regulation, would be unexplained (Oppong et al., 2019).

One of the main components of self-regulated learning is metacognition. Self-regulated learning is the process that involves setting goals, using strategies to achieve goals, monitoring performance until goals are achieved, and evaluating the task (Cogliano et al., 2021). Metacognitive and self-regulatory learning theories attempt to answer how the most effective learning is possible. Acquiring new knowledge in a new domain requires extra cognitive effort, leaving few cognitive resources for monitoring. Self-regulatory learning differs from metacognition only in the fact that it emphasises the motivational and emotional aspects of the person in regulating cognition (Özçakır Sümen, 2021). According to Efklides and Misailidi (2019), metacognition is an important and multifaceted component of self-regulatory learning. A crucial criterion distinguishing the two concepts is that self-regulated learning refers to the learning process itself, whereas metacognition is the outcome and measurement of the learning process. In self-regulated learning, obstacles and errors may arise that may affect the accuracy of metacognition and assessment. Assessment problems arise in situations where students under- or over-estimate their understanding (Coglinao et al., 2021).

1.3.3. The Relationship Between Metacognition and Academic Performance

Several research studies have highlighted the positive relationship between metacognitive processes and academic performance, with metacognition being a significant predictor of academic achievement (Hassan et al, 2022; Jansen et al., 2020; Muncer et al., 2021; Souhila, 2022; Xue et al., 2021). Higher academic achievers have more declarative, procedural and conditioned knowledge than their lower-achieving peers. They are more effective at identifying their cognitive strengths and weaknesses, and the repertoire of learning strategies they use is broad and varied, and they apply

regulatory strategies that are the most appropriate to the learning context (Abdelrahman, 2020; Cai et al, 2019; Concina, 2019). The use of metacognitive learning strategies has a positive impact on academic performance (Smith et al., 2020).

In their meta-analysis, Ohtani and Hisasaka (2018) investigated the relationship between metacognition, intelligence and academic performance, based on the results of 118 studies. Metacognition was moderately correlated with academic performance and intelligence. Both the relationship between metacognition and academic performance and the relationship between metacognition and intelligence were moderated by the method of measuring metacognition, with online methods showing higher effect sizes compared to off-line methods. Metacognition predicted academic performance, after controlling intelligence.

In a meta-analysis, de Boer et al. (2018) examined the long-term effects of teaching learning strategies to improve metacognition on academic performance. From post-test to follow-up measurement, there was a small but significant increase in the effect of strategy instruction on academic performance, suggesting that improving metacognition leads to even better academic outcomes in the long run than in the short run.

1.3.4. The Relationship between Metacognition and Extracurricular Activities

Extracurricular activities develop students' self-regulated learning, thereby leading to positive competence outcomes. Extracurricular activities provide children with supportive, caring adults and mentors, and involve skill-building activities that also develop metacognitive skills (Anwarudin et al., 2021; Bayındır et al., 2021; Feraco et al., 2022; McCosker et al., 2021).

Research on the relationship between competition and metacognitive processes, self-regulatory learning, has concluded that participation in competitions is a significant predictor of metacognition (Anwarudin et al., 2021; Feraco et al., 2022).

Anwarudin et al. (2021) investigated the extent to which students participating in the Mathematics Olympiad apply their metacognitive skills. The students were able to accurately identify the knowledge that was essential to solve the problems, accurately identify the information management strategies used when solving different types of mathematics problems, and check and evaluate the solution of each problem. Similarly, Tohir (2019) concluded in his research that students actively use metacognitive strategies while solving mathematics Olympiad problems.

1.3.5. Developing Metacognition in Education

Metacognition, self-regulated learning and metacognitive awareness are prerequisites for acquiring competences at school and university. It is a means of acquiring literacy, reading and basic mathematical skills in primary school, and higher cognitive processes and operations in secondary school and university. Deficits in metacognitive processes and self-regulated learning may explain underperformance in education (Callan et al., 2020), and students may under- or overestimate their performance (Callender et al., 2016).

1.3.5.1. The development of students' meta-comprehension accuracy

Some research has focused on the development of metacognitive judgements, meta-comprehension accuracy. The research of Callender et al. (2016) aimed at developing students' metacognitive judgements, as preliminary surveys showed that students make incorrect metacognitive judgements immediately after an examination, e.g. under- or overestimating their performance. In their research, they compared students' performance judgements with actual exam performance after 2 exams, with all students being presented with a cognitive bias of overconfidence before the first

exam. After the first exam, despite prior information, the previous metacognitive judgement pattern appeared: low performers overestimated their performance, high performers underestimated their performance. After the first exam, one group of students received feedback on their performance, while the other group did not. After the second exam, both metacognitive accuracy and performance increased in the group that received feedback, but the same improvement was not observed in the control group. This research highlights the essential role of feedback in improving metacognitive accuracy.

Carpenter et al. (2019) also improved students' metacognitive accuracy over 8 sessions using a feedback method. In order to generalise the development of metacognitive accuracy to several different tasks, they used a perceptual discrimination task in the experiment and then examined how metacognitive accuracy developed for this task, later also examining the accuracy of metacognitive judgements on performance in a recognition task independent of the experiment. In the group that received feedback on metacognitive judgements, there was a significant increase in metacognitive accuracy compared to the active control group, where they only received feedback on their performance in the perceptual discrimination task. The improvement shown in the experimental group was generalisable to the independent, non-practiced task, the recognition task as well.

1.3.5.2. Visual representation methods for the development of metacognition

Some researchers have used graphical representation methods to target the development of students' metacognition (Powell et al., 2021; Stevenson et al., 2017).

Powell et al. (2021) developed metacognitive knowledge and meta-comprehension accuracy of pharmacy students using the concept mapping method. Before classes they could read about 14 diseases and during class, they had to make a concept map in groups. On 4 occasions they received feedback on their concept maps, on 3 occasions they did not receive feedback and on 7 occasions they did not have to participate in a group activity on disease processing. Before the classroom discussion of knowledge, to assess meta-comprehension accuracy, students had to estimate how many questions they would be able to answer in the reading comprehension quiz and then they filled in the quiz to test their real comprehension. There was no significant difference in students' metacognitive performance, calculated as the difference between the judged and real quiz performance, across the 3 conditions, but the use of the concept map significantly increased students' performance compared to the days when no concept map was produced.

Stevenson et al. (2017) conducted a meta-analysis to investigate the impact of computer-based concept mapping techniques on the development of self-regulated learning. Concept mapping techniques have positive effects on all three aspects of self-regulated learning, cognitive, metacognitive and motivational strategies. Feedback on performance increases the effectiveness of the method. Technology-based concept mapping methods are more effective in improving self-regulated learning than traditional paper-and-pencil concept maps.

CHAPTER II. RESEARCH OBJECTIVES AND OVERALL METHODOLOGY

The improvement of critical thinking and metacognition as 21st century skills is very important in higher education since these cognitive processes are significant predictors of learning performance, and influence job success (Abueita et al., 2022; Affuso et al., 2022; Hafeez et al., 2022; Ibrahim et al. 2021; Ng et al., 2022; Nikander

et al., 2022). Earlier studies demonstrated that concept mapping elaboration is a very effective visual representation method for the improvement of students' critical thinking skills and dispositions (Aein & Aliakbari, 2017; Carvalho et al., 2020; Huang et al., 2017; Lee et al., 2016; Moattari et al., 2014; Mohammadi et al., 2019; Tseng, 2019; Wang & Liao, 2014). However, the results of some studies suggested that the concept mapping method was not more effective than the traditional teaching method for the improvement of each critical thinking skill and disposition of students (Carvalho et al., 2020; Huang et al., 2012; Mohammadi et al., 2019; Samawi, 2006). Based on these contradictory results of the literature and the limitations of the earlier conducted meta-analyses (Romanko, 2016; Yue et al., 2017), **the first main objective** of this thesis was to investigate the effectiveness of the concept mapping method for the improvement of students' critical thinking skills and dispositions, and to compare the effectiveness of this visual representation method with the traditional teaching method (**Study 1**). More precisely, in our meta-analysis, we also examined the impact of the concept mapping method on subcomponents of critical thinking skills and dispositions. The moderating effect of gender, age, university profile, educational level, concept mapping method characteristics (number of concept maps created by the participants; concept mapping elaboration methods: individual vs. collaborative, paper-pencil vs. digital), and allocation type (the presence of randomisation) were also investigated.

There is a lack of validated instruments in Hungarian language for the measurement of students' metacognition and critical thinking. Based on this absence, **the second main aim** of this thesis was the validation of the Metacognitive Awareness Inventory into Hungarian (MAI) (Harrison & Vallin, 2018; Moxon, 2022; Schraw, & Dennison, 1994) (**Study 2 A**) which is an extensively applied instrument for the measurement of metacognitive awareness, and the Cornell Critical Thinking Test Level Z (CCTT Level Z) (Alias et al., 2022; Ennis et al., 2005 Imperio et al., 2020 Leach et al., 2020) (**Study 2 B**) which is also a frequently used test for the assessment of college students' critical thinking skills. Due to the lack of results regarding the factorial structure of the CCTT Level Z, the main objective of Study 2 A was the comparison of several factorial models (correlated and hierarchical) translated into Hungarian language. Similarly, based on the contradictory results regarding the factorial structure of the MAI, the main objective of Study 2 B was the investigation of the validity of the 52-item version of MAI proposed by Schraw and Dennison and the 19-item version of MAI proposed by Harrison and Vallin on a sample of Hungarian native language students. Our further aim in the Study 2 was the invariance testing of the validated instruments across genders.

Academic excellence is frequently defined as high learning performance. Nevertheless, academic excellence could also manifest itself in the participation in extracurricular activities and the achieved success in high level competitions or Olympiads for instance. Several studies highlighted that the attendance in competitions, Olympiads is positively associated with a high level of critical thinking skills and metacognition (Bayındır et al., 2021; Chen et al., 2020; Feraco et al., 2022; Jatmiko et al., 2020; McCosker et al., 2021; Schiefer et al., 2020; Utomo, 2022). However, earlier studies did not test the differences in metacognition and critical thinking between students attending different types and levels of competitions and did not draw a conclusion regarding what type of competitions students need to participate in to improve their cognitive skills. Based on these above mentioned non-investigated research questions in the literature, **the third main objective** of this doctoral thesis

was to complete the existing literature with results regarding the differences in cognitive skills between students attending and not attending competitions, respectively the main and interaction effects of the participation of students in different types (science subject, humanities and social sciences, sports, art) and levels (school-level competition, county Olympiad and national Olympiad) of competitions on their critical thinking and metacognition (**Study 3**).

Earlier studies demonstrated a positive relationship between the attendance in extracurricular activities and learning performance, indicating that these activities have positive effects on learning outcomes (Feraco et al., 2022; Fujiyama et al., 2021; Kravchenko & Nygård, 2022; Lang, 2021). Similarly, critical thinking is also a significant predictor of learning performance (Abueita et al., 2022; Ali & Awan, 2021; Shahzadi et al., 2020; Xhomara, 2022). However, the moderating effect of competition level, competition type and conference attendance on the association between critical thinking and learning achievement was not assessed in earlier studies. Research regarding the effect of gender on the relationship between critical thinking and learning outcomes indicated contradictory results (Darmaji et al., 2022; Nwuba et al., 2022; Purba, 2022; Wahyudiati, 2022). There is a reduced number of studies regarding the association between critical thinking and learning results examining the difference between students with different academic disciplines (Iqbal et al., 2021; Shahzadi et al., 2020). Based on the findings and the limitations of the literature **the fourth main objective** of the thesis was to investigate the predictive effect of critical thinking on students' learning achievement and the moderating effect of competition level, competition type, conference attendance, academic discipline, and gender on the relationship between critical thinking and learning achievement (**Study 4**).

Several studies confirmed that the concept mapping method improves students' metacognitive processes (Powell et al., 2021; Stevenson et al., 2017), critical thinking (Barta et al., 2022; Carvalho et al., 2020; Khrais & Saleh, 2020; Roshangar et al., 2020; Silva et al., 2022), and learning achievement (Appaw et al., 2021; Ayimbila & Akantagriwon, 2021; Bakolis et al., 2021; Manzon, 2021). Nevertheless, in the case of Psychology and Special Education undergraduate students, or in the case of Hungarian speaking Romanian college students, the effectiveness of the concept mapping method for the improvement of their metacognition and critical thinking was not tested in earlier studies. **The fifth main aim** of this thesis was to examine the effect of the concept mapping method compared with the traditional learning method (rereading) for the improvement of Psychology and Special Education students' critical thinking and meta-comprehension accuracy regarding scientific psychological texts (**Study 5**). The further aim of Study 5 was the investigation of the concept mapping method's effectiveness on students' psychological scientific text comprehension and to test the effect of giving feedback on concept maps on critical thinking, meta-comprehension accuracy and text comprehension achievement.

CHAPTER III. ORIGINAL RESEARCH

3.1. Study 1. The Development of Students' Critical Thinking Abilities and Dispositions Through the Concept Mapping Learning Method – A Meta-Analysis[1]

[1] This study has been published. The current version represents an abbreviated adaptation of the published manuscript.

Barta, A., Fodor, L. A., Tamas, B., Szamoskozi, I. (2022). The development of students critical thinking abilities and dispositions through the concept mapping learning method – A meta-analysis. *Educational Research Review, 37*, 100481. https://doi.org/10.1016/j.edurev.2022.100481

3.1.1. Introduction

The goal of education is to prepare students for life, not only to impart lexical knowledge but to develop 21st century skills that lead to success on the labour market (Chu et al., 2017). Instead of traditional, frontal education, the use of alternative methods is recommended, thanks to which lexical knowledge is transformed into conditional knowledge, making knowledge transfer possible, facilitating the use of knowledge for everyday solutions, thus developing 21st century skills (analytical thinking, metacognition, critical thinking, problem-solving, collaborative communication, creativity skills) (Saleh, 2019; Yennita & Zukmadini, 2021). In traditional teaching, the teacher transmits the information in lecture form, students are mostly passive participants, often taking notes, while group discussion or exercises are rare during the lesson. As a result of this inactivity, students' attention is easily distracted, they do not deeply process the course material, neither do they make connections between the information presented, nor draw conclusions based on the main ideas (Lo & Hew, 2020; Oderinu et al., 2020). Graphical methods used to represent concepts as opposed to traditional education, such as the concept mapping method, allow the graphical mapping of information, the identification of linear and cross-relations in the curriculum, thus contributing to deeper understanding and better academic performance (Chiou et al., 2020; Hwang et al., 2020; Machado & Carvalho, 2020). However, in addition to academic performance, the concept mapping method also has a positive effect on the development of higher-order cognitive abilities, such as metacognition (Prinz et al., 2020; Stevenson et al., 2017) and critical thinking (Carvalho et al., 2020; Chen & Hwang, 2020; Khrais & Saleh, 2020; Mohammadi et al., 2019; Roshangar et al., 2020; Tseng, 2019; Yue et al., 2017), which contribute not only to academic success but also to success in life and career (Strods & Strode, 2018). The concept mapping method is becoming more prevalent in education, being suitable for developing both critical thinking ability (Carvalho et al., 2020; Khrais & Saleh, 2020; Mohammadi et al., 2019; Roshangar et al., 2020; Tseng, 2019) and disposition (Lee et al., 2016; Moattari et al., 2014).

Preliminary reviews and meta-analyses have profile-specifically investigated the effect of the concept mapping method on critical thinking. In the included studies only nursing (Romanko, 2016; Yue et al., 2017) or medical (Daley & Torre, 2010) or speech-language pathology education students (Mok et al., 2008) were analysed. In contrast, we did not exclude research based on university profile but we treated study profile and academic level (secondary school and university students) as moderator variables. The two preliminary meta-analyses (Romanko, 2016; Yue et al., 2017) highlight the positive effect of concept mapping on both critical thinking ability and disposition, but we do not have information on the extent to which the method is effective in developing critical thinking subcomponents. The literature shows great variety in terms of what subcomponents belong to critical thinking ability (Davies & Barnett, 2015; Ennis, 2018; Facione, 1990; Haber, 2020; Halpern & Sternberg, 2020) and what motivational factors and personality traits form the critical thinking disposition (Cui et al., 2021; Davies & Barnett, 2015; Haber, 2020). Yue et al. (2017) presented in their meta-analysis results for different measuring tools, there was no summarisation of results for the same subcomponents measured with different instruments. Thus, in our meta-analysis, we do not only draw a general conclusion

about the effect of concept mapping on critical thinking ability and disposition but also synthesize the results on subcomponents, examining which of these are affected by the concept mapping method as opposed to traditional education. We explore the potential moderating effect of demographic variables on the effectiveness of the method, such as gender and age. We also tested whether the methods of concept mapping (number of concept maps constructed, collaborative elaboration, computerised method) moderate the effect on critical thinking. Preliminary meta-analyses did not take into account the different applications of the method (Romanko, 2016; Yue et al., 2017). In his meta-analysis, Romanko (2016) also included research that did not involve randomisation, whereas Yue et al. (2017) analysed RCTs only. Since randomisation is difficult in many cases in education, quasi-experiments are used instead of randomised research to explore the effect of the independent variable (McMillan & Schumacher, 2014). Similarly, a significant proportion of research on this topic did not use random assignment, so in addition to randomised research, we also included non-randomised research in our analysis, examining how randomisation moderates the effect size (ES).

3.1.2. Research Questions

Our primary research question is whether the concept mapping method is more effective in developing critical thinking skills and dispositions than the traditional educational method. More precisely, in our research, we also examine the impact of concept mapping on subcomponents of critical thinking skills and dispositions. We hypothesise that the concept mapping method, due to its metacognitive process activating function, has a greater impact on complex critical thinking abilities (Davies & Barnett, 2015), and on the cognitive and motivational components of critical thinking disposition (Cui et al., 2021), than the traditional teaching method. Furthermore, we hypothesise that the personality components within critical thinking dispositions (Cui et al., 2021), due to their stable trait nature, are not developed more significantly by the concept mapping method than by the traditional teaching method.

Based on the preliminary research, as well as the weaknesses of the meta-analyses and the unanswered questions, we set up the following additional research questions in which we highlighted the potential effects of several moderators on the effectiveness of the concept mapping method:

1. Do the effects of the concept mapping method differ depending on gender?
2. Does the age of the students affect the effectiveness of the method?
3. Does the effect of the concept mapping method differ depending on the university profile?
3. Does the effect of the concept mapping method differ depending on the level of education? Is the method more effective in the case of secondary school students or university students?
4. Does the number of concept maps affect the ES? Is the effect of development greater if students create more concept maps?
5. Does the concept mapping elaboration method affect the ES? Does collaborative creation have a higher development potential than an individual one?
6. Does the ES differ depending on the digitalisation of concept maps? Does a concept map created on a computerised or online platform have a greater impact than the traditional paper-and-pencil type?
7. Does the effect of the concept map differ depending on the allocation type (random or non-random)?

3.1.3. Method

3.1.3.1. Identification and selection of studies

We conducted a literature search in the Cochrane, Medline, ProQuest, PsycInfo, PubMed, Science Direct, Scopus, Web of Science electronic databases, using the following keywords: 'university students', 'college students', 'undergraduate students', 'secondary school students', 'high school students', 'concept mapping', 'critical thinking', 'critical thinking ability', 'critical thinking disposition'. We also searched for other relevant studies from the reference list of the included articles, previous reviews and meta-analyses.

The studies had to meet the following criteria to be included in the meta-analysis:

1. To include a comparison between an experimental group applying the concept mapping method and a control group participating in traditional education
2. To include critical thinking abilities and/or dispositions as outcomes
3. An evidence-based standardised tool was used to assess critical thinking (for instance we excluded studies that measured critical thinking with the complexity of concept maps or other performance indicators)
4. Quantitative research (studies using a qualitative approach were excluded)
5. Studies with sufficient reported data to calculate the ES
6. Studies published in peer-reviewed journals and studies in English

The literature search identified 3791 studies, the additional search of the reference list of the previous relevant literature revealed 12 additional studies, 247 duplicates were excluded. From the remaining 3556 records we excluded 3418 based on titles and abstracts and examined the full-text of 138 articles. The application of the inclusion criteria reduced the number of studies to 21.

3.1.3.2. Data extraction

We extracted quantitative data, descriptive statistics of the experimental and control group (i.e. means, standard deviations, sample sizes) of each study and each outcome. We also extracted a series of categorical and continuous variables from the included studies for further moderator analysis. Participants' gender (percent of female students), mean age, presence of the randomisation process, students' academic level (secondary school or university) and profile (nursing or English as a foreign language), number of created concept maps, type of concept mapping elaboration method (individual or collaborative; computerised or paper-based) were extracted from the methods section of studies.

Two independent raters evaluated the extracted data from all 21 studies and disagreements were discussed until consensus was reached. Two studies did not report sufficient data to calculate the ES, lacking the SD and the number of participants from each group. The authors of these studies were contacted to provide missing SD and sample size information, but they did not yield the missing data, hence these studies were excluded (Zubaidah et al., 2018).

3.1.3.3. Statistical analysis and interpretation

We executed all analyses using the Comprehensive Meta-Analysis software (CMA, version 2.0). We calculated the individual Hedges' g ESs, measuring the standardised mean difference (SMD) between the concept mapping and traditional learning group. Most often, the SMD was computed utilising means, standard deviations and sample sizes. Where means and standard deviations were not available, we calculated the SMD from other statistics, such as t-values or F values for differences between groups. When a study reported multiple measurement data from the same outcome, the average ES was calculated (Borenstein et al., 2009). A positive Hedges' g reflects that the concept mapping group has higher scores of critical thinking

than the traditional learning group. ESs were interpreted applying Cohen's guidelines, where g = 0.2 indicates small, g = 0.5 medium and g = 0.8 large ESs (Cohen, 1992).

We applied a random-effects model for the calculation of mean ESs due to the heterogeneity of study results. In contrast with the fixed effects model, which assumes that the reason for variability is only the sampling error, the random-effects model assumes the heterogeneity of ESs as the consequence of study design and sampling error (Borenstein et al., 2009). As a result of the variability of the critical thinking outcomes, we grouped them into critical thinking ability and disposition outcome. We conducted separate analyses for these two groups of outcomes. Additionally, we calculated the ESs for each subcomponent of critical thinking ability (analysis, deduction, evaluation, explanation, induction, inference, interpretation, recognition of assumptions) and disposition (analyticity, inquisitiveness, open-mindedness, self-confidence, systematicity, truth-seeking). We conducted a sensitivity analysis excluding outliers which were defined as studies in which the pooled ES's 95% CI (confidence interval) was outside of the 95% CI of the pooled ES (on both sides). 95% CIs were also calculated to determine statistical significance (if the 95% CI does not include zero, this denotes that there is a significant difference between the mean critical thinking outcomes of the concept mapping and the traditional learning group).

Heterogeneity was assessed with I^2 statistic, interpreting 25% as low, 50% as moderate, and 75% as high heterogeneity (Higgins et al., 2003). The potential impact of categorical moderator variables was measured by conducting subgroup analysis using the mixed-effects model, which consists of a fixed-effects procedure across subgroups and a random-effects procedure within subgroups (Borenstein et al., 2009). We applied meta-regression with a restricted maximum likelihood model for continuous moderator variables (Borenstein et al., 2009).

We investigated small study effects with several methods. Publication bias was analysed through the visual inspection of the funnel plot, which helps in distinguishing publication bias from other factors of asymmetry (Peters et al., 2008). We also employed the Egger's test for the asymmetry of the funnel plot (Egger et al., 1997) and the trim and fill method (Duval & Tweedie, 2000) as complementary procedures in investigating potential publication bias or small study effects.

3.1.4. Results

The 21 selected studies included 108 relevant comparisons between the concept mapping and the traditional learning groups, with 872 participants in the concept mapping (experimental) group and 823 in the traditional learning (control) group.

3.1.4.1. The effect of the concept mapping method compared to the traditional learning method on critical thinking ability outcomes

For the critical thinking ability outcome 19 studies were pooled, with an ES of g = 0.531, 95% CI 0.279 to 0.783, with substantial heterogeneity (I^2 = 82%). Exclusion of four potential outliers led to a small decrease in ES, g = 0.490, 95% CI 0.312 to 0.668, and reduced heterogeneity (I^2 = 55%). With regard to critical thinking subcomponents, we found significant moderate ESs for deduction ability (g = 0.661, 95% CI 0.124 to 1.197), with considerable heterogeneity (I^2 = 91%), and for inference (g = 0.578, 95% CI 0.036 to 1.120), with substantial heterogeneity (I^2 = 90%). We also detected significant but smaller ESs for induction (g = 0.493, 95% CI 0.134 to 0.852), with moderate heterogeneity (I^2 = 67%), and small ES for recognition of assumptions (g = 0.337, 95% CI 0.111 to 0.562), without heterogeneity (I^2 = 0%). For analysis (g = 0.638, 95% CI -0.041 to 1.317, I^2 = 92%), evaluation (g = 0.196, 95% CI -0.087 to 0.479, I^2 = 75%), explanation (g = 0.155, 95% CI -0.577 to 0.887, I^2 = 87%), and

interpretation (g = 0.131, 95% CI -0.213 to 0.474, I^2 = 63%) no statistically significant differences were found. Follow-up outcomes were only reported in two studies for critical thinking ability.

3.1.4.2. The effect of the concept mapping method compared to the traditional learning method on critical thinking disposition outcomes

For the critical thinking disposition outcome 5 studies were pooled, g = 0.648, 95% CI 0.266 to 1.031, with moderate heterogeneity (I^2 = 66%). For the critical thinking disposition outcome no outliers were identified. A significant high ES for truth-seeking was demonstrated, g = 0.994, 95% CI 0.227 to 1.761, with high heterogeneity (I^2 = 86%). A significant moderate ES for analyticity was found g = 0.753, 95% CI 0.204 to 1.301, with considerable heterogeneity (I^2 = 75%). Similarly, moderate ES was detected for inquisitiveness, g = 0.591, 95% CI 0.342 to 0.840, and for open-mindedness, g = 0.568, 95% CI 0.320 to 0.817, without heterogeneity (I^2 = 0%). The ES is non-significant for self-confidence, g = 0.373, 95% CI -0.100 to 0.846, with moderate heterogeneity (I^2 = 65%), and for systematicity, g = 0.492, 95% CI -0.115 to 1.099, with considerable heterogeneity (I^2 = 80%). In the included studies follow-up outcomes were not reported for critical thinking disposition.

3.1.4.3. The moderating effect of students' demographic characteristics

Univariate meta-regression indicated a non-significant relationship between students' gender as a continuous variable (percent of females) and critical thinking ability ES (slope = -0.008, 95% CI: -0.027 to 0.010). After the exclusion of the outliers, the relationship remained statistically non-significant (slope = -0.001, 95% CI: -0.012 to 0.008).

Students' gender was significantly associated with critical thinking disposition ES (slope = -0.018, 95% CI: -0.034 to -0.002); an increase in the percentage of female participants leads to a decrease in critical thinking disposition.

We found a non-significant relationship between students' mean age and critical thinking ability ES (slope = 0.027, 95% CI: -0.051 to 0.105); even with the exclusion of outliers this result remains non-significant (slope = 0.027, 95% CI: -0.005 to 0.060).

The association between mean age and critical thinking disposition could not be tested because the mean age of the participants was provided in only one study out of five on critical thinking disposition.

3.1.4.4. The moderating effect of educational conditions

Students' educational level was not a significant moderator of ES for critical thinking ability (p = 0.840). Similarly, academic discipline was not a significant moderator of ES for critical thinking ability (p = 0.109).

We could not test the moderating effect of the educational level and the academic discipline on critical thinking disposition ES due to the lack of studies including primary or secondary school students. Only in one study was the effect of the concept mapping method on English as a foreign language students' critical thinking disposition investigated, in the other studies nursing students participated.

3.1.4.5. The moderating effect of concept mapping elaboration methods

There is a non-significant association between the number of concept maps created by the students and critical thinking ability ES (slope = 0.041, 95% CI: -0.006 to 0.090), after the exclusion of outliers, this result remains non-significant (slope = 0.010, 95% CI: -0.029 to 0.050). Similarly, there is no significant relationship between the number of concept maps and critical thinking disposition ES (slope = 0.004, 95% CI: -0.075 to 0.084).

It was revealed that the collaborative development of concept maps was not a significant moderator of ES for critical thinking ability (p = 0.266), nor the digital method (p = 0.756).

We could not test the moderating effect of the collaborative elaboration and digital method on critical thinking disposition ES due to the lack of studies including these methods. Only one included study applied the collaborative and digital method for the development of students' critical thinking disposition.

3.1.4.6. The moderating effect of allocation type

The allocation type was a significant moderator for the comparison between concept mapping and traditional learning for critical thinking ability (p = 0.043). Effects were moderate for randomised studies (11 trials, g = 0.739, 95% CI 0.356 to 1.122), with high heterogeneity (I^2 = 83%). The ES in non-randomised studies was low (8 trials, g = 0.265, 95% CI 0.014 to 0.517), with moderate heterogeneity (I^2 = 66%).

For critical thinking disposition the allocation sequence was not a significant moderator of the ES (p = 0.860).

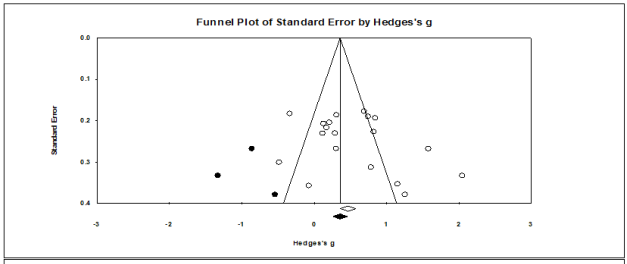
3.1.4.7. Small study effects and publication bias

Visual inspection pointed to an asymmetrical funnel for both critical thinking ability and disposition (see Figure 1). Egger's regression intercept test was not statistically significant for critical thinking ability (intercept = 3.174, 95% CI: -1.748 to 8.097, p = 0.191) and critical thinking disposition (intercept = 4.525, 95% CI: -2.243 to 11.295, p = 0.123), thus revealing no evidence for publication bias/small study effects.

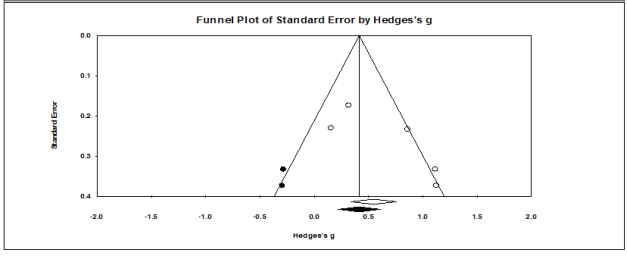
Figure 1

Funnel plot for comparison between the concept mapping and the traditional learning method for critical thinking ability (A) and disposition (B) outcomes

A



B



The Duval & Tweedie's trim and fill method revealed three potentially missing studies for critical thinking ability, which, if included would lead to the decrease of the ES from 0.531 (95% CI: 0.278 to 0.783) to 0.353 (95% CI: 0.073 to 0.633). For critical thinking disposition the trim and fill method estimated two potentially missing studies which, if imputed, would lead to the decrease of ES from 0.648 (95% CI: 0.265 to 1.031) to 0.426 (95% CI: 0.041 to 0.810).

3.1.5. Discussion and Conclusions

In the present meta-analysis, we found a moderate effect of the concept mapping method on students' critical thinking ability and disposition improvement compared with the traditional education method. The effect of concept mapping differs in diverse critical thinking ability and disposition subcomponents, it is more useful for the development of critical thinking abilities, such as recognition of assumptions, induction, deduction, inference and dispositions, like analyticity, inquisitiveness, open-mindedness, truth-seeking, than the traditional teaching method. These results highlight the superiority of the use of the concept mapping method in the improvement of critical thinking abilities and dispositions as compared to the traditional learning method. Through the representation of concepts and their linear and cross relationships (Novak, 2010), students have more accurate inductive and deductive reasoning skills, reflect relevant information more precisely, recognise assumptions more accurately than students using traditional learning methods.

We analysed the impact of different moderators on the effect of concept mapping on critical thinking. Except for the randomisation process (for critical thinking ability) and gender (for critical thinking disposition), the tested moderating effects are not significant. We could draw more accurate conclusions regarding the differences in the effectiveness of this method for academic levels and disciplines, respectively the effect of different elaboration types for critical thinking abilities and dispositions, if more studies including the tested subgroups were conducted, reporting additional data about students' characteristics and the description of the method. Many existent studies are poorly reported. Detailed reporting of the statistical data for enhancing knowledge about the applicability of concept mapping would be very useful.

The findings of the meta-analysis draw attention to the effectiveness of the concept mapping method as an active learning, metacognitive regulation strategy for teaching and applying it in educational settings. The deep processing of course material and the exploration of the relationships between relevant concepts presuppose the functioning of metacognitive processes, which are activated by the use of the concept mapping method (Powell et al., 2021; Stevenson et al., 2017). The method can be applied in a variety of ways, either in paper-pencil or digital format, individually or in groups, and the introduction of the method to students is time and energy efficient. Providing feedback on the created concept maps further enhances the effectiveness of the method in promoting cognitive processes (Powell et al., 2021; Stevenson et al., 2017), as well as academic performance (Dmoshinskaia et al., 2021; Joseph et al., 2017). Finally, this meta-analysis compares the effectiveness of the concept mapping method with the traditional, passive one. Further studies could contrast the applicability of concept mapping with other active teaching methods.

3.2. Study 2. A. The Comparison of Several Factorial Structures of the Cornell Critical Thinking Test Level Z[2]

3.2.1. Introduction

Critical thinking, as a 21st century ability, was examined in several psychological areas. Educational psychology research concentrates on the role of critical thinking in learning, teaching and learning achievement, respectively on its improvement in the case of students aswell as in the case of teachers (Renatovna, & Renatovna, 2021; Yuan et al., 2022). Cognitive psychology studies focus on the association between critical thinking and other cognitive processes, and also investigate the role of critical thinking in complex cognitive activities, like problem solving (Işıklar, & Abalı-Öztürk, 2022; Song et al., 2022) and metacognition (Boran, & Karakuş, 2022; Deliligka, & Calfoglou, 2022). In clinical psychology it is mainly the role of the maladaptive form of critical thinking, self-criticism that is studied in several mental disorders (Moroz, & Dunkley, 2019; Wakelin et al., 2022). In personality psychology the relationships between critical thinking dispositions and personality traits (Eshmirzaeva, 2020; Toker, & Akbay, 2022), as well as effective personality (Merma-Molina et al., 2022) were measured.

One of the most frequently applied instruments for the assessment of critical thinking is the Cornell Critical Thinking Test (Alias et al., 2022; Ennis et al., 2005), that is based on the Cornell/Illinois model (Ennis et al., 2005; Imperio et al., 2020). Two versions of the test were developed for the measurement of critical thinking in two different age groups. The Cornell Critical Thinking Test Level X was developed for the measurement of critical thinking of middle school and high school students (4-12 grades), while the Cornell Critical Thinking Test Level Z is applicable for the measurement of advanced and gifted high school students', college students' and adults' critical thinking abilities (Ennis et al., 2005). Both tests have forced- choice question format, dichotomous items. The Cornell Critical Thinking Test Level Z has 52 items, and measures cognitive abilities such as deduction, induction, observation, credibility, meaning, assumption identification (Gunawardena & Wilson, 2021; Ennis et al., 2005). The CCTT is a domain general instrument, the results are independent of domain specific knowledge.

The CCTT Level Z is a widely used instrument for the measurement of students' critical thinking in educational, cognitive psychology studies (Heidari, 2020; Saud, 2020), however, there is a lack of empirical studies in the literature investigating the factorial structure of the CCTT Level Z. The instrument contains seven subtests: deduction, meaning and fallacies, observation and credibility of sources, induction: hypothesis testing, induction: planning experiments, definition and assumption identification. The induction ability is composed of two subscales, the test measures the application of induction in hypothesis testing and in planning experiments. Similarly, the test contains two sections for the measurement of assumption identification ability (Ennis et al., 2005). The test authors emphasise the interdependency and overlap of the measured cognitive abilities. Due to the reduced number of items assessing the separate cognitive abilities, the test authors suggest taking into consideration the total score for drawing conclusions regarding individual

[2] This study is submitted for publication and it is under review. The current version represents an abbreviated adaptation of the submitted manuscript.

Barta, A., Tamás, B., Póka, T. (2023). *The Comparison of Several Factorial Structures of the Cornell Critical Thinking Test Level Z* [Manuscript submitted for publication].

differences in critical thinking. In the manual of the Cornell Critical Thinking Test the reliability results of the instrument regarding undergraduate and graduate student samples are reported. The Kuder-Richardson reliabilities ranged between .50 and .76, and the split-half reliabilities werebetween .49 and .80 (Ennis et al., 2005; Verburgh et al., 2013).

3.2.2. The Aims of the Study

The main aim of the study was the translation into Hungarian language of the first instrument that measures high school students', gifted students' and adults' critical thinking abilities.Due to the lack of earlier studies investigating the factorial structure of the Cornell Critical Thinking Test Level Z, based on the recommendations of the authors of the test (Ennis et al., 2005), beside the unidimensional model, our aim was the testing and comparison of several correlated and hierarchical factorial models (two-factor models: deductive reasoning, inductive reasoning; three-factor models: deduction, induction, meaning and fallacies; four- factor models: deduction, induction, meaning and fallacies, assumption identification) translated into Hungarian language. Our further aim was the invariance testing of the CCTT across genders.

3.2.3. Methods

3.2.3.1. Participants

825 Hungarian-speaking students of Babeș-Bolyai University participated in the study. There were no multiple outliers identified, based on the Mahalanobis distance. 78.3% of the participants were females, with a mean age of 21.76 years (SD = 7.12). The youngest person was 18, and the oldest was 64 years old. The majority of participants were first-year (85%), 13.9% of them were third year and only 1.1% of the participants were second-year undergraduate students. 86.7% of the participants were full-time students, 42.2% of them studied at the Faculty of Psychology and Educational Sciences.

3.2.3.2. Instrument

The Cornell Critical Thinking Test Level Z was applied for the measurement of students' higher order thinking abilities (Ennis et al., 2005). The CCTT contains 52 items and seven sections: deduction (1-10), meaning and fallacies (11-21), observation and credibility of sources (22-25), induction (Hypothesis Testing) (26-38), induction (Planning Experiments) (39-42), definition and assumption identification (43-46), assumption identification (47-52). The CCTT is a multiple-choice test, has a forced-choice question format, and dichotomously scored items, the answer for an item can be correct or incorrect. The retranslation method wasapplied for the translation of the instrument into Hungarian language, executed by two experts. The participation in the study was preceded by the informed consent of participants, the test was completed in Google Forms.

3.2.3.3. Data analysis

For the CFA analysis we applied the Mplus version 8.7 with weighted least squares, mean and variance adjusted estimation (WLSMV; Muthén et al., 1997; Muthén & Muthén, 2021). The following absolute fit indices were used to evaluate model-data fit: Chi-squared, Root Mean Squared Error of Approximation (RMSEA), Standardised Root Mean Square Residual (SRMR). Some relative fit indices, like Comparative Fit Index (CFI), and Tucker Lewis Index (TLI) were also applied. Acceptable model-data fit criteria included CFI > 0.90, TLI > 0.90, RMSEA < 0.05, SRMR < 0.08 (Awang, 2012; Hu & Bentler, 1999; Kline, 2015), while models having CFI and TLI values greater than .95 indicate excellent fit (Hu & Bentler, 1999; Kline, 2015). The relative fit of the alternative factor structures was compared based onthe

difference between CFI values. Post hoc inspection was conducted on the best fitting model, the following item retention criteria was applied: items with significant factor loadings and items with factor loadings ≥ 0.32 (Costello & Osborne, 2005; Leach et al., 2020).

We also tested measurement invariance across genders with Muthén's (2013) two-step procedure for dichotomous data (Leach et al., 2020). We have tested configural and scalar invariance models because metric invariance testing is not allowed for binary variables. The configural model was applied to determine if the four-factor structure existed across the two groups, without the application of equality constraints. In the scalar invariance model, the factor loadings and intercepts were constrained as equal across groups (Leach et al., 2020; Muthén, 2013; Muthén & Muthén, 2017). Due to our relatively large sample size and the sensitivity of the absolute $\Delta\chi 2$ to sample size, ΔRMSEA and ΔSRMR was used for testingthe change in model fit, values < .015 and < .01 indicated no difference between the models (Chen, 2007; Meade et al., 2008).

3.2.4. Results

Based on the descriptive statistic results, item means varied substantially with values ranging from .11 (Item 18) to .80 (Item 8), and standard deviations ranged from .30 (Item 18) to .50 (Items 4 and 29). These values reveal that several items were very difficult (e.g., Items 12, 18, 32, 37), while other items were easier (e.g., Items 2, 8, 17, 26, 46).

3.2.4.1. The comparison of several factorial structures of the CCTT Level Z in Hungarian

Based on the recommendations of the test authors (Ennis et al., 2005), beside the unidimensional model, several correlated and hierarchical models [(two-factor models: deductive reasoning (Items 1-25, 43-52), inductive reasoning (Items 26-42); three-factor models: deduction (Items 1-10, 43-52), meaning and fallacies (Items 11-21), induction (Items 26-42); four-factor models: deduction (Items 1-10), meaning and fallacies (Items 11-21), induction (Items 26-42), assumptions (43-52)] were also tested. The analysis of the observation and credibility of sources as a discrete factor and the items belonging to this factor were excluded from the CFA analysis due to the insignificant factor loadings of allfour items and negative covariances, residual variances with other latent variables. Due to the overparameterised factor structure, negative covariances, residual variances, the items belonging to the induction-hypothesis testing and induction-planning experiments subscales were treated as the induction factor, similarly, the definition-assumption identification and assumption identification subscales were also treated as a single factor.

The CFA results of the tested models revealed that the four-factor correlated [$\chi 2$ (1068) = 1247.415, CFI = .909, TLI = .904, RMSEA = .014, SRMR = .061] and the four-factor secondorder models [$\chi 2$ (1070) = 1248.971, CFI = .909, TLI = .904, RMSEA = .014, SRMR = .061] exceeded the minimum criteria for acceptable model fit on the majority of fit indices (Awang,2012; Hu & Bentler, 1999; Kline, 2015).

3.2.4.2. The abbreviated Hungarian version of the CCTT Level Z

Due to the very similar statistical fit indices of the two four-factor models, for post hoc inspection the second-order four factor model was chosen based on the theoretical approach of the test developers, who suggest the interpretation of critical thinking as a constructinvolving subfactors (Ennis et al., 2005). Based on the post hoc inspection of the four-factor second-order model, items loaded insignificantly to the deduction factor (3, 4) were not retained. Similarly, three items were identified (13,

15, 19) that were not loaded significantly to the meaning and fallacies factor, two items with insignificant factor loadings to theinduction factor (39, 40), and one item (43) that loaded insignificantly to the assumption identification factor. Within the deduction subfactor five items were identified, (1, 2, 6, 9, 10) within the meaning and fallacies subfactor five items (12, 14, 16, 20, 21), within the induction subfactor five items (31, 35, 38, 41, 42), and within the assumptions subfactor four items (44, 45, 51, 52) with factor loadings < .32. After the exclusion of items with insignificant and lower than .32 factor loadings (Costello, & Osborne, 2005; Leach et al., 2020), the abbreviated version of the test included 22 items, three items (5, 7, 8) loaded significantly to the deduction factor, four items (11, 17, 18, 20) to the meaning and fallacies factor, ten items (26, 27, 28, 29, 30, 32, 33, 34, 36, 37) to the induction factor and five items (46, 47, 48, 49, 50) to the assumption identification factor. The abbreviated 22-item four- factor second-order model indicated excellent fit indices [$\chi 2$ (203) = 259.309, CFI = .967, TLI = .963, RMSEA = .018, SRMR = .056].

The internal consistency of the 22-item Hungarian version of the test was acceptable (α = .601).

3.2.4.3. Measurement invariance across genders of the Hungarian version of the CCTT level Z

Measurement invariance across genders (males and females) of the 22-item abbreviated version of the Cornell Critical Thinking Test Level Z was tested using configural and scalar models. The models showed acceptable model fit indices for both genders at every level of invariance. The configural model indicated the 22-item four-factor structure for males and females. The scalar model indicated that each of the item loadings onto the factors and item intercepts are similar across gender groups. The changes in fit indices (ΔCFI and ΔRMSEA) indicated no significant differences in relative model fit between the configural and the restrictive model, which confirm the configural and scalar invariance of the instrument across genders.

3.2.5. Discussion and Conclusions

The CCTT is an extensively used instrument for the measurement of critical thinking in different psychological areas, but there is a lack of earlier studies in the literature investigating its factorial structure and reliability. Due to the questionable factorial structure of the test and the absence of instruments assessing adults' critical thinking abilities, the aim of the current study was the translation of the test into Hungarian language and the analysis of different factorial structures proposed by the test developers (Ennis et al., 2005). Comparing the unidimensional model with different correlated and hierarchical factorial structures, our results revealed that the four-factor structure of the test, including four higher order cognitive abilities within the general critical thinking factor, namely the deduction, meaning and fallacies, induction and assumption identification factors, is the most applicable, reliable and valid model for the measurement of critical thinking of undergraduate Hungarian-speaking students from Babeș-Bolyai University. Post hoc inspection of the four-factor second-order structure indicated a 22-item shortened version of the test with excellent fit indices.

The internal consistency of the shortened version of the test is similar to the findings of test authors (Ennis et al., 2005), who examined the CCTT's internal consistency in the case of several undergraduate and graduate samples, revealing Kuder-Richardson reliabilities between .50 and .76. The current internal consistency indicator (.601) exceeds the reliability of the original version of CCTT applied in the

study of Verburgh et al. (2013), who found a lower Cronbach alpha indicator (.52) in the case of Belgian Educational Science students.

Additionally, invariance testing of the instrument across genders was conducted. The results revealed configural and scalar invariance across genders of the 22-item four-factor structure Hungarian version of the CCTT, confirming that the test has similar factorial structure, item loadings and intercepts in the case of male and female groups.

Based on the limitations of the study, it is recommended for future research to evaluate the factorial structure and reliability of the instrument on older master students, on samples with balanced gender, academic discipline, academic year distribution, replacing the convenience sampling method with the cluster one. Similarly, it would be beneficial to translate and validate the test into other languages, its measurement invariance analysis between groups formed based on socio-demographic and academic characteristics of students, in order to complete the literature with additional results regarding the factorial structure and the reliability of the test.

This is the first study that aimed at the comparison of different factorial structures of the CCTT level Z. Similarly, there is a lack of instruments in Hungarian language measuring students' and adults' critical thinking. The results of this study, the translated and validated Hungarian version of the CCTT level Z, are very important for the empirical measurement of critical thinking skills in education, and in different work environments. The empirical measurement of students' critical thinking provides information about their level of thinking skills and promotes the identification of those skills that require further development in educational context.

3.3. Study 2. B. Validation and Evaluation of the Factorial Structure of Two Versions of the Metacognitive Awareness Inventory in Hungarian[3]

3.3.1. Introduction

Metacognitive processes were investigated within several psychological areas (Norman et al., 2019). Numerous earlier studies demonstrated a significant positive relationship between metacognitive processes and learning achievement (Cai et al., 2019; Hassan et al., 2022; Jansen et al., 2020; Muncer et al., 2021; Souhila, 2022; Xue et al., 2021). However, several research results established only weak association between students' metacognition and learning performance (Abu Bakar, & Ismail, 2020; Zhao et al., 2019), or did not find any significant correlation (Vettori et al., 2018), or detected a negative association (Hashmi et al., 2019). One possible explanation for these contradictory research results could be the low validity and reliability of instruments measuring metacognition.

The measurement of metacognition is difficult because it cannot be defined as an explicit behaviour, but neither as a whole implicit process, because we are aware of our metacognitive processes to a certain degree. Metacognitive processes could not be observed in a direct manner. It is a complex mental process that contains metacognitive knowledge and regulation, but also includes momentary motivational and emotional aspects that could affect the control of cognitive processes (Lai, 2011).

[3] This study is submitted for publication and it is under review. The current version represents an abbreviated adaptation of the submitted manuscript.

Barta, A., Tamás, B., Póka, T. (2023). *Validation and Evaluation of the Factorial Structure of Two Versions of the Metacognitive Awareness Inventory in Hungarian* [Manuscript submitted for publication].

The 52-item Metacognitive Awareness Inventory (MAI) was developed by Schraw and Dennison in 1994. The inventory contains two subscales: metacognitive knowledge and metacognitive regulation. Metacognitive knowledge includes three subcomponents: declarative, procedural, and conditional knowledge, while the five subcomponents within metacognitive regulation are: planning, comprehension monitoring, information management strategies, debugging strategies and evaluation.

The two-factor structure of MAI was confirmed in several studies, but the items belonging to the factors vary across research. Harrison and Vallin (2018) have tested the factorial structure of MAI on American undergraduate population, applying confirmatory factor analysis (CFA) and multidimensional random coefficients multinomial logit item-response modelling. They confirmed the two-factor structure, but from the original 52 items only 19 items loaded significantly on the two factors. In 2020, Rao and Jaiswal tested the MAI factor structure on Indian high school students. The results of the EFA showed that from the 52 items only 25 items were related to the two factors, to metacognitive knowledge and regulation, that was also confirmed by a CFA. The instrument was denominated as the MAI short version, the two factors significantly associated to learning achievement. Moxon (2022) aimed at the confirmation of the two-factor structure of the MAI proposed by Harrison and Vallin (2018) on Japanese population. Based on the CFA results, the model fitted only the minimally acceptable criteria. After the post hoc inspection, the 13-item abbreviated two-factor model had good fit indices. The two-factor structure of MAI was also confirmed on Malaysian secondary school students, respectively measurement invariance across age groups demonstrated the same two-factor structure in the two age groups (16-17 years, and 18-19 years students) (Siang & Lan, 2011).

Craig and his colleagues (2020) analysed in a systematic review the results of studies investigating the factorial structure of instruments measuring metacognition. From the included 22 studies, 12 studies tested the factorial structure of MAI with EFAs and CFAs. From the 22 studies 13 confirmed the presence of two factors: metacognitive knowledge and metacognitive regulation. The results regarding the subfactors of the two main factors show a high variety. In the case of MAI, the different studies identified 3, 4, 5, 6 and 8 subcomponents. The eight-factor structure of MAI proposed by Schraw and Dennison was confirmed only on its Persian (Pour & Ghanizadeh, 2017) and Turkish (Akin et al., 2007) versions. Within the eight factors three belonged to the metacognitive knowledge and five to the metacognitive regulation factor. Therefore, these studies also confirm the two-factor structure. The factorial structure of MAI was also tested on Indian medical and dental first year college students (Omprakash et al., 2021). The EFA results demonstrated six factors (declarative knowledge, procedural knowledge, conditional knowledge, planning, evaluation, monitoring), after the exclusion of 12 items. Based on the CFA results of Teo and Lee (2012) on Asian students, the eight-factor structure was not supportable, a three-factor 21-item structure was proposed.

Due to the diversity of the results a precise conclusion regarding the subcomponents of MAI proposed by Schraw and Dennison cannot be drawn. Proving the two-factor structure and the items belonging to these factors failed in numerous studies. Despite the contradictory results, based on the findings of Craig and his colleagues' (2020) systematic review and meta-analysis, it could be concluded that the two-factor structure of MAI is the best acceptable.

3.3.2. The Aims of the Study

As a consequence of the contradictory results regarding the factorial structure of the MAI, the main aim of the study is the investigation of the applicability and validity of the 52-item version of MAI proposed by Schraw and Dennison and the 19-item version of MAI proposed by Harrison and Vallin on a sample of Hungarian native language students. Our goal is to find out which of the two mentioned versions of MAI provides more reliable results regarding metacognitive knowledge and regulation. Beside testing the two factorial structures, our further aim is the investigation of the internal consistency of the instrument, respectively invariance testing across genders for the examination of the stability of the factorial structure and internal consistency of the instrument, independently of genders.

3.3.3. Methods

3.3.3.1. Participants

Originally, 819 undergraduate students from Babeş-Bolyai University participated in the study, however, after the exclusion of multiple outliers identified with the Mahalanobis distance, the final sample was formed from 770 students. The majority of participants were females (79.2%), with a mean age of 21.84 years (SD = 7.27), the youngest student was 18 years old, the oldest participant's age was 64 years. The majority of the participants were first-year students (84.4%), 14.5% of participants were in their third, and only 1% of them were in their second academic year. 86.1% of the participants were full-time students, 42.9% of students studied at the Faculty of Psychology and Educational Sciences.

3.3.3.2. Instruments

The Metacognitive Awareness Inventory developed by Schraw and Dennison (1994) was applied. The MAI contains 52 items that can be grouped in one of eight subscales: three subscales within metacognitive knowledge (declarative, procedural and conditional knowledge), five subscales within metacognitive regulation (planning, information management strategies, debugging strategies, comprehension monitoring and evaluation). Participants responded on a 5-point Likert scale, where = Not at all typical of me, 2 = Not very typical of me, 3 = Somewhat typical of me, 4 = Fairly typical of me, and 5 = Very typical of me (Harrison & Vallin, 2018; Moxon, 2022). This scale differs from the original, semantic-differential response scale, that contains yes-no type, dichotomous answers. The Likert scale indicates a more detailed picture regarding the degree of metacognitive awareness but could also increase the acquiescence bias (Krosnick & Presser, 2010). For the translation of the instrument into Hungarian language the retranslation method was applied. The participation in the study was preceded by the informed consent of participants, the questionnaire was completed through Google Forms.

3.3.3.3. Data analysis

We tested the two-factor model of MAI proposed by Schraw and Dennison in 1994, but due to the results of earlier studies on different samples that could not replicate the 52-item structure of MAI (Harrison, & Vallin, 2018; Moxon, 2022; Omprakash et al., 2021; Rao, & Jaiswal 2020; Teo, & Lee, 2012), we also tested the 19-item version of MAI proposed by Harrison and Vallin (2018).

Outliers were identified and excluded based on Mahalanobis d^2. For the assessment of internal consistency of the tested models Cronbach alphas, correlations between factors, AVE and CR scores were calculated.

For the CFA analysis we used SPSS Amos version 24 (Arbuckle, 2016) with maximum likelihood estimation. The chi-squared absolute fit index is sensitive to sample size (Vandenberg, 2006), researchers have proposed alternative fit indices for

the assessment of model fit (Kline, 2015; Tabachnick & Fidell, 2007). The following absolute fit indices were used to evaluate model-data fit: Chi-squared, Goodness of fit index (GFI), Root Mean Squared Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR). Some relative fit indices: Comparative Fit Index (CFI), Tucker Lewis Index (TLI), and the $\chi 2/df$ as the parsimonious fit index were also applied. Acceptable model-data fit criteria included GFI > .90, CFI > 0.90, TLI > 0.90, RMSEA < 0.05, SRMR < 0.08, $\chi 2/df < 5$ (Awang, 2012; Hu & Bentler, 1999; Kline, 2015), while models having GFI, CFI and TLI values greater than .95 indicate excellent fit (Hu & Bentler, 1999; Kline, 2015). For the comparison of the goodness of fit of the models Akaike and Bayesian information criteria were also used, lower values indicating better data fit.

Measurement invariance across males and females of the MAI was assessed with a series of invariance models. Configural, metric, scalar and residual invariance models were tested by sequentially adding restrictions on parameters, applying a stepwise approach. The configural model was used to determine whether the two-factor structure existed across the two groups, without the application of equality constraints. In the metric model the configural testing was used with the additional constraint of factor item loadings equal between groups, to analyse the item factor loadings equivalence across the two groups. Metric testing was applied for the scalar model, in this case item intercepts were also constrained equal between groups to examine the equivalence of item intercepts across groups. The scalar model was applied for the residual invariance testing, constraining item residuals equal between groups to analyse whether the residuals of the MAI items are similar across gender groups (Putnick & Bornstein, 2016). Due to the sensitivity of the absolute $\Delta\chi 2$ to sample size, descriptive fit indices were used, ΔRMSEA and ΔSRMR, for testing the change in model fit, values < .015 indicated no difference between less restrictive and more restrictive models (Chen, 2007; Meade et al., 2008).

3.3.4. Results

3.3.4.1. The 52-item version of Metacognitive Awareness Inventory

Based on the Mahalanobis d^2 49 outliers were identified. The Cronbach alpha indicators revealed good internal consistency of the 52-item version of MAI for the metacognitive knowledge (α = .804) and metacognitive regulation (α = .875) subscales. The calculated CR were also acceptable, for metacognitive knowledge it was .797, for metacognitive regulation .872. However, the calculated AVE scores were very low, for the metacognitive knowledge subscale it was 0.198 and for the metacognitive regulation subscale it was .171. The correlation between the two subscales was high, r (768) = .674, p < .001. The CFA analyses of this first tested, 52-item model based on Schraw and Dennison's theory did not satisfy all of the criteria for acceptable model fit (Hu & Bentler, 1999), $\chi 2$ (1192) = 2712.068, $p < .001$, GFI = .871; RMSEA = .041, 90% CI [.039, .043], SRMR = .052; CFI = .835, TLI = .817; $\chi 2/df = 2.275$. Information criteria for the 52-item model were AIC = 3084.068, BIC = 3948.296.

3.3.4.2. The 19-item version of Metacognitive Awareness Inventory

The Cronbach alphas of the 19-item version of MAI indicated adequate internal consistency for the metacognitive knowledge (α = .713) and metacognitive regulation (α = .739) subscales. The calculated CR were also acceptable, for metacognitive knowledge it was .735, for metacognitive regulation .735, but the calculated AVE scores were very low, for the metacognitive knowledge subscale was .275 and for the metacognitive regulation subscale it was .203. The correlation between the two

subscales was high, r (768) = .555, p < .001. The CFA results of the second tested Harrison-Vallin 19-item model showed that every evaluated fit index exceeded the minimum criteria for acceptable model fit (Awang, 2012; Hu & Bentler, 1999; Kline, 2015), and well-fitting model criteria on the majority of fit indices, excepting the $\chi 2$, CFI and TLI values (Hu & Bentler, 1999; Kline, 2015), $\chi 2$ (137) = 298.885, $p < .001$, GFI = .961; RMSEA = .039, 90% CI [.033, .045], SRMR = .04; CFI = .934, TLI = .918; $\chi 2/df = 2.182$. Information criteria for the 19-item model were AIC = 404.885, BIC = 651.144, factor loadings are presented in Table 1.

Table 1

Standardised and unstandardised regression weights of the 19 – item Harrison-Vallin model

Item	Subcategory	Standardised Estimate	Unstandardised Estimate	Unstandardised Standard error
10	K (DK)	.331	1.000	
16	K (DK)	.396	1.195	.161
20	K (DK)	.394	1.161	.172
32	K (DK)	.348	.954	.137
27	K (PK)	.693	2.342	.289
33	K (PK)	.660	2.163	.269
26	K (CK)	.459	1.607	.228
35	K (CK)	.726	2.356	.289
6	R (P)	.456	1.000	
8	R (P)	.475	1.083	.123
39	R (IMS)	.360	.672	.094
41	R (IMS)	.430	.927	.115
43	R (IMS)	.446	.978	.116
21	R (M)	.505	1.099	.121
24	R (E)	.522	1.298	.140
50	R (E)	.489	1.213	.136
40	R (DS)	.423	.850	.105
44	R (DS)	.341	.648	.094
51	R (DS)	.478	.905	.102

Notes. K = metacognitive knowledge, R = metacognitive regulation, DK = declarative knowledge, PK = procedural knowledge, CK = conditional knowledge, P = planning, IMS = information management strategies, M = monitoring, DS = debugging strategies, E = evaluation.

3.3.4.3. Measurement invariance of the 19-item version of Metacognitive Awareness Inventory

Measurement invariance across genders (males and females) of the 19-item Harrison-Vallin model was tested using a series of multi-group CFAs, configural, metric, scalar, and residual models. All tested models showed acceptable model fit indices for both genders at every level of invariance. The configural model indicated the 19-item two-factor structure for males and females. The metric model indicated that each of the item loadings is similar across gender groups, the scalar model suggested that each of the item intercepts are similar across groups and the residual model revealed that the item residuals for each of the 19 items are also similar across groups.

The descriptive fit indices (ΔRMSEA and ΔSRMR) indicated no significant differences in relative model fit between less restrictive and more restrictive models,

which confirm the configural, metric, scalar and residual invariance of the instrument across genders.

The Cronbach alphas of the 19-item version of MAI indicated adequate internal consistency for male ($\alpha_{\text{metacognitive knowledge}} = .723$, $\alpha_{\text{metacognitive regulation}} = .736$) and female responders ($\alpha_{\text{metacognitive knowledge}} = .735$, $\alpha_{\text{metacognitive regulation}} = .736$). The calculated CR were also acceptable for males (for metacognitive knowledge it was .730 and for metacognitive regulation .730) and for females (for metacognitive knowledge it was .739 and for metacognitive regulation .728). The calculated AVE scores were very low in the case of the male (for the metacognitive knowledge subscale it was .271 and for the metacognitive regulation subscale it was .206) and the female (for the metacognitive knowledge subscale it was .278 and for the metacognitive regulation subscale it was .442) groups. The correlation between the two subscales in the male group was moderate, $r(158) = .441$, $p < .001$, in the female group it was high, $r(608) = .527$, $p < .001$.

3.3.5. Discussion and Conclusions

The MAI is an extensively applied instrument for the measurement of metacognition in different areas of psychology research, but the originally recommended factor and item structure was reproduced only in a few studies, which questions the reliability of the instrument and the accuracy and applicability of the results of the inventory. Because of the contradictory literature results and the absence of instruments for the measurement of metacognition validated to Hungarian language, in the present study the MAI was translated into Hungarian language with the retranslation method.

The main aim of the study was the investigation and comparison of the factorial structure and internal consistency of the original 52-item version of MAI proposed by Schraw and Dennison, and the 19-item version of MAI, developed by Harrison and Vallin with CFA and multidimensional random coefficients multinomial logit item-response modelling methods, on Hungarian native language undergraduate student population.

The internal consistency of both models of MAI based on Cronbach alpha indicators and calculated CR (composite reliability) showed adequate internal consistency for both subscales (metacognitive knowledge and regulation). However, the calculated AVE (average variance extracted) scores were under the acceptable criteria in the case of both versions for each subscale. The correlation between the two subscales of the 52-item and 19-item version MAI was high.

Based on the CFA results the 52-item Schraw and Dennison model indicated unacceptable GFI, CFI and TLI values (Hu & Bentler, 1999), but the 19-item Harrison-Vallin model indicated acceptable and excellent model fit indices on every evaluated criteria (Awang, 2012; Hu & Bentler, 1999; Kline, 2015). Correspondingly, the 19-item model indicated better AIC and BIC values. Similarly to the majority of earlier studies (Craig et al., 2020; Harrison & Vallin, 2018; Moxon, 2022; Ning, 2016; Rao & Jaiswal, 2020; Siang & Lan, 2011), which investigated the factorial structure of the MAI, the two-factor structure of MAI was also confirmed on Hungarian native language undergraduate student population. Based on our results, the 19-item two-factor version of MAI, proposed by Harrison and Vallin is reliably applicable on the Hungarian native language undergraduate students of Babeș-Bolyai University, furthermore, the instrument is appropriate for drawing conclusions regarding metacognitive knowledge and regulation in the studied population.

Measurement invariance across genders was tested on our better fitted 19-item Harrison-Vallin model, applying a series of multigroup-CFAs. The four tested models (configural, metric, scalar, and residual) indicated acceptable model fit indices for both gender groups. The configural model supported the 19-item two-factor structure for the group of males and females. The metric model indicated that the item loadings are similar, the scalar model showed that the item intercepts are similar, and the residual model demonstrated that the item residuals are similar across the two groups. The descriptive fit indices (ΔRMSEA and ΔSRMR) indicated measurement invariance (configural, metric, scalar and residual) across the two gender groups.

The internal consistency of the 19-item MAI across genders was also tested. The Cronbach alpha values and the calculated CR (composite reliability) scores indicated adequate internal consistency, but the calculated AVE (average variance extracted) values were very low. The correlation between the metacognitive knowledge and regulation subscales for males was moderate and for females it was high. Based on the invariance testing results, the 19-item two-factor structure is reliably applicable on Hungarian native language undergraduate student population, independently of gender.

Due to the limitations of the study, for future studies it is recommended to analyse the factorial structure and reliability of the instrument on older, other native language academic population, with balanced gender and Faculty, academic profile distribution, that could be achieved by replacing the convenience sampling method with stratified random sampling method.

3.4. Study 3. Excellence-Based Differences of Undergraduate Students in Metacognition and Critical Thinking

3.4.1. Introduction

Critical thinking and metacognition can be improved with the application of learning and teaching techniques, such as methods usable for visual representation of knowledge that aim the activation of complex thinking processes (Barta et al., 2022; Powell et al., 2021). On the other hand, numerous earlier studies raised attention to the positive effect of the participation in extra- and co-curricular activities (competitions, conferences, group memberships) on students' higher order cognitive abilities (Feraco et al., 2022; Jatmiko et al., 2020; McCosker et al., 2021; Schiefer et al., 2020).

Academic excellence is frequently defined as the high learning achievement of students, but participation in extracurricular activities, like school-level competitions, county-level, national-level Olympiads, competition in conferences during the academic years, also improve the knowledge and expertise within a certain scientific, art, or sports field (Abueita et al., 2022; Ibrahim et al., 2021; Rebholz et al., 2022).

The participation in extracurricular scientific, art or sports activities, in competitions, conferences, and scientific group membership, are not only significant predictors of learning achievement (Abueita et al., 2022; Akpur, 2020; Ali & Awan, 2021; Ibrahim et al., 2021; Ng et al., 2022; Shahzadi et al., 2020), but several studies also demonstrated the positive effect of these activities on students' critical thinking, divergent thinking, reasoning ability (Chen et al., 2020; Huang & Yeh, 2017; Jatmiko et al., 2020; Schiefer et al., 2020).

In some studies students' critical thinking skills were improved with techniques that were based on the combination of coopetition, cooperative methods, and competitive situations (Huang & Yeh, 2017; Wang et al., 2017). The critical thinking skills and dispositions of journalism students were also developed with the combination of cooperation and competition, accomplished through an online

gamified platform (Huang & Yeh, 2017). Utomo et al. (2022) compared the critical thinking skills of regular, accelerated and Olympiad class students. Olympiad class students interpreted the questions correctly, applied accurate problem-solving strategies, and with the help of these correctly selected methods drew precise conclusions. The preparation process for the Olympiads, the proficiency in Olympiad tasks, the experiences and expertise acquired in the competitions, promoted the outstanding level of critical thinking skills.

Participation in natural science activities, competitions, Olympiads are significant predictors of students' critical thinking (Chen et al., 2020; Jatmiko et al., 2020; Nugroho et al., 2019; Rusdin & Rusli, 2020; Utomo, 2018; Wahidin & Romli, 2020). The solution of mathematical problems, tasks, and mathematics competition attendance have a positive association with the level of critical thinking (Rusdin & Rusli, 2020; Utomo, 2018; Wahidin & Romli, 2020). Similarly, students attending chemistry and robotics competitions have higher critical thinking skills than their non-attending peers (Chen et al., 2020; Jatmiko et al., 2020; Nugroho et al., 2019).

Participation in extracurricular, scientific, art or sports activities improves students' field specific knowledge, their knowledge regarding their own abilities and regulation skills, self-regulated learning, and metacognition (Anwarudin et al., 2021; Bayındır et al., 2021; Feraco et al., 2022; McCosker et al., 2021; Tohir, 2019). Students attending competitions acquire effective time management skills, learn to differentiate between the easier and more difficult tasks during the intensive preparation process, in this manner they can precisely identify the amount of time required by the tasks with different difficulty (Salmeen et al., 2019). Through the preparation process they obtain expertise and reach conclusions regarding the information strategies required by the tasks, thereby they acquire more metacognitive knowledge than their peers who do not attend competitions (Anwarudin et al., 2021; Tohir, 2019). International Olympiad winner athletes reported that their experience acquired in the competitions promoted the improvement of their planning ability and self-control (Jordalen et al., 2019).

Numerous earlier studies established a positive association between mathematics achievement and metacognition (Desoete & De Craene, 2019; Ohtani & Hisasaka, 2019). The solution of mathematics tasks requires effective information representation, planning, evaluative and analytical skills, so attendance at mathematical competitions or Olympiads improves students' self-regulation (Anwarudin et al., 2021; Tohir, 2019). Combinatorics courses improved students' metacognition, the application of effective task solving strategies, and the optimal use of the conclusions drawn from the evaluation of their own achievement (Tohir & Muhasshanah, 2021). Idiege and Nja (2021) confirmed the effectiveness of an interactive competition-based metacognitive strategy to increase the learning achievement of chemistry students.

3.4.2. Aims of the Study and Hypotheses

Due to the reduced number of previous studies in the literature investigating competition-based variables' effect on higher order cognitive skills, the main aim of the study is to complete the existing literature with results regarding the differences in critical thinking and metacognitive awareness (metacognitive knowledge and regulation) between students attending or not attending competitions. Our further aim is to explore the differences in critical thinking and metacognitive awareness (metacognitive knowledge and regulation) between students attending higher level (Olympiads) and lower level (school-level) competitions, students attending

competitions in science subjects, humanities and social sciences, sports and arts, additionally, to investigate the interaction effect of the competition level (no competition attendance, school-level competition, Olympiad attendance) and competition type (science subject, humanities and social sciences, sports, art) on students' critical thinking and metacognition.

Based on the results of earlier studies and on the objectives of the study the following hypotheses were created:

1. There is a significant effect of the competition level and the competition type on students' critical thinking ability.
2. There is a significant effect of the competition level and the competition type on students' metacognitive awareness (metacognitive knowledge and metacognitive regulation).

3.4.3. Methods

3.4.3.1. Participants

579 undergraduate students of Babeș-Bolyai University participated in this study. Most participants were women (78.6%), first year (83.8%), full-time students (86.4%) of the Faculty of Psychology and Educational Sciences (40.8%). 24.5% of the students didn't participate in any competition, 30.7% of the students attended school-level competitions, 29.7% participated in county Olympiads and 15% in national Olympiads. The distribution of students based on competition type was the following: 21.1% participated in science subject, 38.3% in humanities or social science, 9% in sports and 7.1% in art competitions. The mean age of participants was 21.78 years (SD = 7.14). The lowest score at critical thinking was 4, the highest was 19, and the mean score was 11.77 (SD = 3.24). The minimum score at metacognitive knowledge was 14, the maximum was 40, the minimum score at metacognitive regulation was 20, the maximum was 54. The mean score of metacognitive knowledge was 27.91 (SD = 4.15) and that of metacognitive regulation was 39.29 (SD = 5.35).

3.4.3.2. Instruments

3.4.3.2.1. Demographic questionnaire

The demographic questionnaire was developed for the measurement of socio-demographic characteristics of students (gender, age, study type – full time or distance learning, academic year, academic discipline), participation in competitions, competition level, and competition type. In the case of participation in competitions the students selected the highest level of competition in which they participated (non-attendance, school level competition, county level Olympiad, national Olympiad), and the type of the highest-level competition that they attended (science subject, humanities or social sciences, sports, art).

3.4.3.2.2. The Hungarian version of the Cornell Critical Thinking Test Level Z

The test has forced-choice question format, contains 22 items and four sections: deduction (5, 7, 8), meaning and fallacies (11, 17, 18, 20), induction (26, 27, 28, 29, 30, 32, 33, 34, 36, 37), assumption identification (46, 47, 48, 49, 50). Based on the recommendations of the original test developers (Ennis et al., 2005), and due the reduced and unequal number of items of some subscales, the further analysis regarding critical thinking was conducted using the total score of the students' critical thinking.

3.4.3.2.3. The Hungarian version of Metacognitive Awareness Inventory

The instrument includes 19 items, and two subscales: metacognitive knowledge (10, 16, 20, 26, 27, 32, 33, 35) and metacognitive regulation (6, 8, 21, 24, 39, 40, 41, 43, 44, 50, 51) (Harrison & Vallin, 2018; Moxon, 2022). A 5-point Likert scale was used (1 = Not at all typical of me, 2 = Not very typical of me, 3 = Somewhat typical

of me, 4 = Fairly typical of me, and 5 = Very typical of me) (Harrison & Vallin, 2018; Moxon, 2022). The instrument showed adequate internal consistency for both subscales ($\alpha_{\text{metacognitive knowledge}} = .722$, $\alpha_{\text{metacognitive regulation}} = .706$).

3.4.3.3. Research design

An ex post facto design was applied, investigating the effect of the earlier participation of students in different types and levels of competitions on their critical thinking and metacognition. The compared groups were formed based on students' participation in competitions, competition level and types of competition. The effect of these quasi-experimental variables was tested on the dependent variables: metacognitive awareness (metacognitive knowledge and regulation) and critical thinking.

3.4.3.4. Data analysis

A priori power analysis via G*Power3 (Faul et al., 2007) for two-way (3x4) ANOVA based on type I error with a p-value of .05 and statistical power of 0.80 with twelve groups showed that for a medium effect size ($f = .25$) the required sample size is n = 225, for two-way (3x4) MANOVA for a medium effect size ($f^2(V) = .0625$) the required sample size is n = 192. For the investigation of the effect of competition level and competition type on students' critical thinking two-way (3x4) ANOVA was conducted. For the inspection of the effect of competition level and competition type on students' metacognitive awareness (metacognitive knowledge and regulation) two-way (3x4) MANOVA was applied. Based on the results of the two-way MANOVA, the effect of competition level on students' metacognitive awareness was further analysed, splitting the Olympiad group into two subgroups: county and national Olympiad groups.

3.4.4. Results

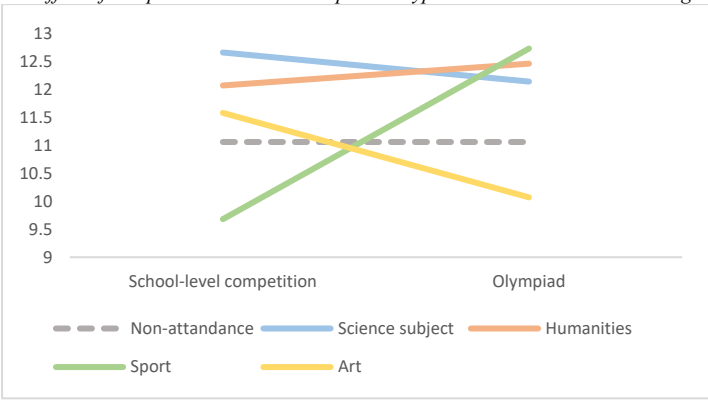
3.4.4.1. The effect of competition level and competition type on students' critical thinking

A 3 x 4 ANOVA test was conducted to assess the effect of competition level (no competition attendance, school-level competition, Olympiad attendance) and competition type (science subject, humanities, sports, art) on students' critical thinking. Two outliers in the science subject Olympiad group and one outlier in the art competition group were detected; these three cases were excluded from further analysis. Levene's test indicated that the assumption of homogeneity of variances was met, $F_{\text{Levene}}(8, 567) = 1.274$, $p = .254$.

The main effect of competition level was not statistically significant, F (1, 567) = .706, $p = .401$, $\eta_p^2 = .001$. The main effect of competition type was statistically significant, F (3, 567) = 3.440, $p = .017$, $\eta_p^2 = .018$, indicating that competition type accounted for a 1.8% of the variance in critical thinking. The interaction effect also was significant, F (3, 567) = 3.836, $p = .010$, $\eta_p^2 = .020$, indicating that 2% of the variance in critical thinking was explained by the level and the type of the competition.

Figure 1

The effect of competition level and competition type on students' critical thinking



Scheffe post hoc comparisons were conducted to evaluate pairwise differences between group means because of the unequal sample size of the groups (Nawi et al., 2020). The pairwise comparisons between the means of the competition level groups revealed a significant difference in critical thinking between students who didn't participate in any competitions and students who participated in Olympiads ($p = .001$), indicating that in the Olympiad attending group ($M = 12.26$, $SD = 3.088$) the critical thinking scores of students were significantly higher than the scores in the non-attending group ($M = 11.06$, $SD = 3.169$). There was no statistically significant difference between students attending school-level competitions ($M = 11.63$, $SD = 3.410$) and Olympiads ($p = .124$), nor between students not attending competitions and students attending school-level competitions ($p = .275$). The post hoc test regarding the competition type groups indicated that in the non-attending group ($M = 11.06$, $SD = 3.169$) there were lower scores at critical thinking than in the science subject competition group ($M = 12.38$, $SD = 3.331$, $p = .023$) and the humanities competition group ($M = 12.37$, $SD = 3.011$, $p = .005$). Students who attended sports competitions ($M = 10.33$, $SD = 3.508$) had lower scores than students who attended science subject ($p = .004$) and humanities ($p = .002$) competitions. There was no statistically significant difference between the non-attending and sports competition ($p = .719$), non-attending and art competition ($M = 11.05$, $SD = 2.978$, $p = 1.000$), science subject and humanities competition ($p = 1.000$), science subject and art competition ($p = .253$), humanities and art competition ($p = .204$), sports and art competition ($p = .879$) groups (see Figure 1).

3.4.4.2. The effect of competition level and competition type on students' metacognitive awareness

A 3 x 4 MANOVA test was conducted to examine the effect of competition level (no competition attendance, school-level competition, Olympiad attendance) and competition type (science subject, humanities, sport, art) on students' metacognitive awareness. Metacognitive awareness includes two dependent variables: metacognitive knowledge and metacognitive regulation.

The main effect of competition level was not statistically significant, F (2, 564) = 1.42, $p = .243$, Pillai's Trace = .005, $\eta_p^2 = .005$. Similarly, competition type had no

statistically significant effect on students' metacognition, F (6, 1130) = .644, p = .695, Pillai's Trace = .007, η_p^2 = .003. The interaction effect of competition level and competition type was not statistically significant, F (6, 1130) = .747, p = .612, Pillai's Trace = .008, η_p^2 = .004.

Based on the scores of metacognitive knowledge an increasing tendency was detected in all Olympiad groups. Therefore, the effect of the level of competition was further analysed splitting the Olympiad group into two new subgroups: county and national Olympiad groups. In the MANOVA test the independent variable was the level of competition (no competition attendance, school-level competition, county Olympiad, national Olympiad attendance) and the dependent variable was the metacognitive awareness (metacognitive knowledge and metacognitive regulation).

The effect of competition level was statistically significant, F (6, 1140) = 2.227, p = .038, Pillai's Trace = .023, η_p^2 = .012. The effect of competition level on metacognitive knowledge is significant, F (3, 570) = 3.811, p = .010, η_p^2 = .020, indicating that 2% of the variance for metacognitive knowledge was explained by the competition level. The level of competition effect on metacognitive regulation was not statistically significant, F (3, 570) = .376, p = .770, η_p^2 = .002.

Scheffe post hoc comparisons indicated a significant difference in metacognitive knowledge between students who had attended national level Olympiads (M = 29.29, SD = 3.677) and school-level competitions (M = 27.58, SD = 4.347, p = .020), between the national level Olympiad and county level Olympiad (M = 27.71, SD = 4.158, p = .040) groups. No statistically significant difference was detected between the non-attending (M = 27.72, SD = 4.043) and the school level competition (p = .993), the non-attending and the county level Olympiad (p = 1.000), the non-attending and the national level Olympiad (p = .052), nor between the school-level competition and the county level competition (p = .993) groups in metacognitive knowledge scores.

The Scheffe post hoc test indicated no statistically significant differences between students attending different levels of competitions in metacognitive regulation scores.

3.4.5. Discussion and Conclusions

The aim of this study was to complete the existing literature regarding the effect of participation in competitions on students' higher order cognitive abilities. Earlier studies didn't analyse the possible differences in metacognition and critical thinking between students attending different levels and types of competitions. The results of the study emphasise the significant effect of students' participation in competitions on their higher order cognitive abilities. More specifically, the level and the type of these competitions have a significant effect on students' critical thinking skills and their metacognitive awareness.

Significant differences were found in critical thinking between students not attending competitions and students attending Olympiads; there was no significant difference between students attending school-level and Olympiad competitions. Similarly, differences between students attending different types of competitions were observed. Students attending science subject and humanities competitions have significantly higher critical thinking scores than students not attending competitions and students who participated in sports competitions. Our results support the results of earlier studies investigating the relationship between participation in competitions, Olympiads and critical thinking, the positive effect of the participation in science subject, like chemistry, robotics, mathematics competitions (Chen et al., 2020; Jatmiko et al., 2020; Nugroho et al., 2019; Rusdin & Rusli, 2020; Wahidin & Romli, 2020).

Competition type had no statistically significant effect on students' metacognitive awareness (metacognitive knowledge and regulation), similarly, students' metacognition was not influenced significantly by competition level when it was operationalised with three levels (non-attendance, school-level competition, and Olympiad attendance), but an ascending tendency of metacognitive knowledge was found with the increase of competition level, regardless of the competition type. The effect of competition level on metacognition was further analysed, splitting the Olympiad attendance into two levels (county and national Olympiad), and a significant effect on metacognitive awareness was found. Within metacognition a statistically significant effect could be found only on students' metacognitive knowledge. Students attending national level Olympiads have higher metacognitive knowledge than students who attended lower-level competitions. The results regarding the effect of participation in competitions on metacognition uphold the results of earlier studies investigating the effect of competitions, Olympiads on students' metacognition and self-regulation (Anwarudin et al., 2021; Bayındır et al., 2021; Feraco et al., 2022; Jordalen et al., 2019; Salmeen et al., 2019; Tohir, 2019). Based on our results, the first hypothesis, regarding the effect of competition level and competition type on students' critical thinking has proved to be true. We can accept the second hypothesis only partly, because the competition level has a significant effect on metacognition, but the main effect of competition type and the interaction effect of competition level and competition type on metacognition was not statistically significant.

In order to increase of the generalisability of our results and the equal representation of students based on their socio-demographic and competition-related variables, the application of the cluster sampling method instead of convenience sampling, the involvement of other-language-speaking students, master and doctoral students, beside Hungarian-speaking undergraduate students, is recommended for future studies. With the help of interrupted time series quasi-experimental design it would be feasible to apply the multiple measurement of cognitive abilities before and after students' participation in different types and levels of competitions, in order to draw more precise conclusions regarding the effect of competitions on critical thinking and metacognition. In the interest of decreasing the response bias, on-line methods for the measurement of metacognition are recommended, beside the applied self-reported instrument.

The results of the study highlight the positive effects of competitions on students' higher order cognitive abilities, the importance of competitions being organised by educational institutions (schools, universities). Teachers and parents have a prominent role in finding the field that students are the most interested in, in the identification of their talents, in supporting their improvement in these fields, and in motivating them to participate in competitions.

3.5. Study 4. The Moderating Effect of Competition Level, Competition Type, Conference Attendance, Gender and Academic Discipline on the Relationship between Critical Thinking and Learning Achievement

3.5.1. Introduction

Earlier studies demonstrated a positive, significant relationship between students' critical thinking and learning performance (Abueita et al., 2022; Akpur, 2020; Ali & Awan, 2021; Shahzadi et al., 2020). Competition attendance also is a significant predictor of academic achievement (Chen et al., 2020; Chen & Chang, 2020; Makhdum et al., 2023). Besides broadening the subject-specific knowledge of students participating in competitions, that contributes in a direct way to their eminent

performance in the current subject as well. The development of critical thinking skills also explains the positive relationship between competition attendance and learning achievement (Akbar et al., 2022; Chiang et al., 2023; Greyling, 2023; Jatmiko et al., 2020; Rif'at et al., 2022). Students solve numerous subject-specific tasks, problems during the competitions, and in the competition preparation period, that contributes to the improvement of their inductive thinking, through the detection of similarities between the different tasks, they identify general problem-solving rules that they could apply in solving similar tasks (Ersteniuk et al., 2020; Zubova et al., 2021). With the help of deductive thinking, they could apply the general rules (for example grammatical rules, mathematical formulas, physical laws) in specific task situations (Rif'at et al., 2022; Yang et al., 2020).

The question occurs whether the level or the type of competition, or participation in a scientific conference increases the strength of association between critical thinking and learning achievement. Earlier studies also examined the effect of gender on the relationship between critical thinking and learning performance, but the results of these studies are contradictory (Darmaji et al., 2022; Iqbal et al., 2021; Nwuba et al., 2022; Purba, 2022). Beside all these, some earlier studies investigated the association between critical thinking and learning achievement in the case of students with different academic disciplines (Ali & Awan, 2021; Ibrahim et al., 2021), and the differences between students learning at different academic majors (Iqbal et al., 2021; Shahzadi et al., 2020), so our study also tries to answer the question if there is a difference in the relationship between critical thinking and learning performance between students learning different academic disciplines.

3.5.2. Aims of the Study and Hypotheses

The main aim of this study was to explore moderator variables which influence the relationship between critical thinking and learning performance measured in baccalaureate results and academic achievement. More specifically, our aim was to investigate the predictive effect of critical thinking on students' learning achievement and the moderating effect of competition level (no competition attendance, school-level competition, county Olympiad, national Olympiad attendance), and of competition type (no competition attendance, science subject, humanities, sports, and art competition attendance) on the relationship between critical thinking and learning achievement. Our further goal was to analyse the effect of participation in conferences on the relationship between critical thinking and learning achievement. Beside the investigation of the effect of these extracurricular activities on students' learning achievement, our aim was to examine the effect of students' academic discipline (social sciences, humanities, natural sciences, sport) and gender on the association between critical thinking and learning achievement.

Based on the results of earlier studies and on the aims of the study the following hypotheses were created:

1. Competition level has a significant predictive effect on students' learning achievement and has a significant moderating effect on the relationship between critical thinking and learning achievement.
2. Competition type has a significant predictive effect on students' learning achievement and has a significant moderating effect on the relationship between critical thinking and learning achievement.
3. Conference attendance has a significant predictive effect on students' learning achievement and has a significant moderating effect on the relationship between critical thinking and learning achievement.

4. Academic discipline has a significant predictive effect on students' learning achievement and has a significant moderating effect on the relationship between critical thinking and learning achievement.
5. Gender has a significant predictive effect on students' learning achievement and has a significant moderating effect on the relationship between critical thinking and learning achievement.

3.5.3. Methods

3.5.3.1. Participants

816 undergraduate students participated in the study. Most participants were women (78.4%), first year (84.8%), full-time students (87.1%), the mean age was 21.71 years (SD = 7.09). 50% of the participants had social sciences, 16.7% humanities, 25.7% natural sciences and 7.6% sports as an academic profile. 24.9% of the students did not participate in any competitions, 30.4% attended school-level competitions, 29.5% participated in county Olympiads and 15.2% in national Olympiads. 20.1% of the students attended humanities/social sciences, 39.5% science subject, 8.8% sports and 6.7% attended art competitions. Only 19.5% of the students participated in conferences.

3.5.3.2. Instruments

A demographic questionnaire was used for measuring students' socio-demographic data (gender, age, study type – full time or distance learning, academic year, academic discipline), mean baccalaureate result, academic achievement, participation in competitions, competition level (school-level competition, county level Olympiad, national Olympiad), competition type (science subject, humanities, sports, art), participation in conferences.

For the assessment of students' critical thinking ability, the Cornell Critical Thinking Test Level Z – Hungarian language abbreviated version was used. The test has forced-choice question format, including 22 items and four subscales: deduction (5, 7, 8), meaning and fallacies (11, 17, 18, 20), induction (26, 27, 28, 29, 30, 32, 33, 34, 36, 37), assumption identification (46, 47, 48, 49, 50). The internal consistency of the test (α = .601) is similar with the internal consistency results of earlier studies applying the original test format (Ennis et al., 2005, Verburgh et al., 2013).

3.5.3.3. Research design

For the evaluation of the moderating effects of competition level, competition type, conference attendance, academic discipline and gender on the association between students' critical thinking ability and learning achievement a correlational, cross-sectional design was used. The predictor variable is critical thinking ability, the outcome variables are baccalaureate and academic results, and the moderator variables are competition level, competition type, conference attendance, academic discipline and gender.

3.5.3.4. Data analysis

A priori power analysis via G*Power3 (Faul et al., 2007) for hierarchical linear regression based on type I error with a p-value of 0.05 and statistical power of 0.80 with four tested predictors (total number of predictors 9) showed that for a medium effect size (f^2 = 0.15) the required sample size is n = 85, while for a small effect size (f^2 = 0.02) the required sample size is n = 602. The moderator analyses were performed using bootstrapping method, with Process macro version 4.0 (Hayes, 2022).

3.5.4. Results

Students' mean baccalaureate result was 8.29 (SD = .87), the mean academic achievement was 9.12 (SD = .70), and their mean critical thinking ability was 11.78 (SD = 3.24).

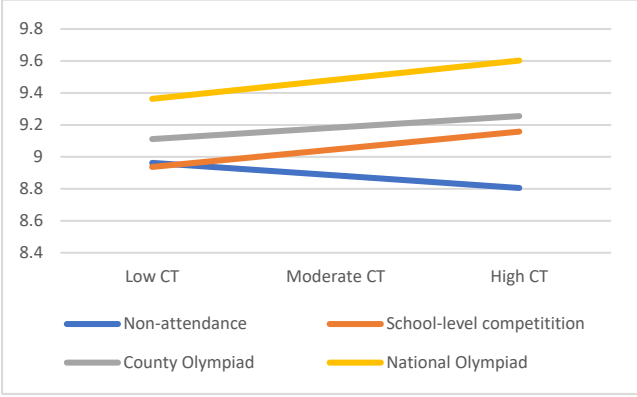
3.5.4.1. Level of competition as a moderator of the association between critical thinking and learning performance

For the investigation of the moderating effect of level of competition on the relationship between students' critical thinking and baccalaureate results a moderator analysis was performed. The overall moderation model is significant, indicating that the competition level and critical thinking accounted for a 14.1% of the variance for baccalaureate results [F (7, 808) = 21.982, $p < .001$, $R^2 = .141$]. Students' critical thinking is a significant predictor of baccalaureate results [$b_1 = .076$, t (808) = 3.652, $p < .001$]. School-level competition attendance does not significantly predict baccalaureate results [$b_2 = .013$, t (808) = .166, $p = .868$], but county Olympiad attendance [$b_3 = .315$, t (808) = 3.823, $p < .001$] and national Olympiad attendance [$b_4 = .564$, t (808) = 5.761, $p < .001$] are significant predictors of baccalaureate results. The interaction effect of critical thinking and school-level competition for baccalaureate results was not statistically significant [$b_5 = -.003$, t (808) = -.110, $p = .913$], similarly there was no significant interaction effect of critical thinking and county Olympiad attendance [$b_6 = -.020$, t (808) = -.760, $p = .447$], respectively of critical thinking and national Olympiad attendance [$b_7 = .002$, t (808) = .055, $p = .956$] for baccalaureate results. It was revealed that level of competition is not a significant moderator of the relationship between critical thinking and baccalaureate results [F (3, 808) = .340, $p = .796$, $\Delta R^2 = .001$].

The moderating effect of level of competition on the relationship between critical thinking and academic mean results was also tested. The overall moderation model is significant, indicating that the competition level and critical thinking accounted for a 9.6 % of the variance for academic achievement [F (7, 808) = 16.932, $p < .001$, $R^2 = .096$]. Students' critical thinking is not a statistically significant predictor of academic results [$b_1 = -.024$, t (808) = -1.556, $p = .120$]. School-level competition attendance is a significant predictor of academic achievement [$b_2 = .164$, t (808) = 2.535, $p = .011$], similarly county Olympiad attendance [$b_3 = .299$, t (808) = 4.450, $p < .001$], respectively national Olympiad attendance [$b_4 = .599$, t (808) = 8.337, $p < .001$] are significant predictors of academic results. The interaction effect of critical thinking and school-level competition for academic mean results was statistically significant [$b_5 = .058$, t (808) = 3.034, $p = .003$]. Similarly, there was a significant interaction effect of critical thinking and county Olympiad attendance [$b_6 = .047$, t (808) = 2.362, $p = .018$], respectively of critical thinking and national Olympiad attendance [$b_7 = .061$, t (808) = 2.755, $p = .006$] on academic achievement. It was revealed that the level of the competition is a significant moderator of the relationship between critical thinking and academic achievement [F (3, 808) = 3.601, $p = .013$, $\Delta R^2 = .012$]. In the case of students who attended school-level competitions, with one unit increase of critical thinking, academic achievement also increases by .034 points [$b = .034$, t (808) = 3.025, $p = .003$], in the case of students who attended national Olympiads one unit increase of critical thinking is associated with the improvement of academic results by .037 points [$b = .037$, t (808) = 2.334, $p = .020$] (Figure 1).

Figure 1

The moderating effect of level of competition on the relationship between critical thinking and academic mean results

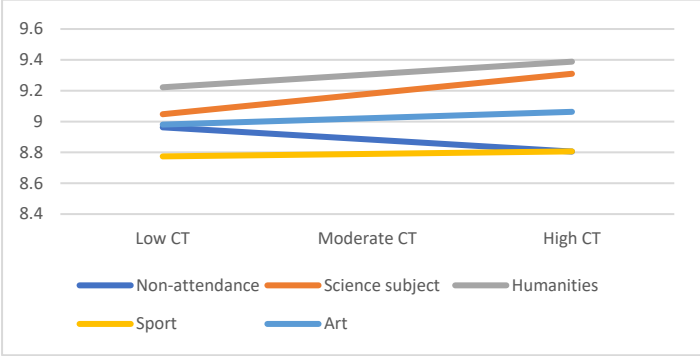


3.5.4.2. Type of competition as a moderator of the association between critical thinking and learning performance

The overall moderation model regarding the moderating effect of type of competition on the relationship between critical thinking and baccalaureate results is significant, indicating that the competition types and critical thinking accounted for a 13.7 % of the variance for baccalaureate grades [$F (9, 806) = 14.663$, $p < .001$, $R^2 = .137$]. Students' critical thinking is a statistically significant predictor of baccalaureate results [$b_1 = .076$, $t (806) = 3.662$, $p < .001$]. Science-subject competition attendance [$b_2 = .267$, $t (806) = 2.918$, $p = .003$], and humanities competition attendance [$b_3 = .378$, $t (806) = 4.802$, $p < .001$] are significant predictors of baccalaureate grades. However, participation in sports [$b_4 = -.162$, $t (806) = -1.355$, $p = .176$] and in art competitions [$b_5 = .004$, $t (806) = .025$, $p = .980$] are not significant predictors of baccalaureate results. The interaction effects of critical thinking and science subject competition attendance [$b_6 = -.035$, $t (806) = -1.293$, $p = .196$], critical thinking and humanities competition attendance [$b_7 = -.012$, $t (806) = -.475$, $p = .635$], critical thinking and sports competition attendance [$b_8 = .008$, $t (806) = .253$, $p = .801$], critical thinking and art competition attendance [$b_9 = .047$, $t (806) = .994$, $p = .320$] were not statistically significant for baccalaureate test results. The moderating effect of type of competition on the association between critical thinking and baccalaureate results is not statistically significant [$F (4, 806) = 1.162$, $p = .326$, $\Delta R^2 = .005$].

Figure 2

The moderating effect of type of competition on the relationship between critical thinking and academic mean results



Investigating the effect of type of competition on the relationship between critical thinking and study mean results, our overall moderation model is significant, competition type and critical thinking accounted for a 9.6 % of the variance for academic grades [F (9, 806) = 10.110, p < .001, R^2 = .096]. Students' critical thinking does not significantly predict their mean study results [b_1 = -.024, t (806) = -1.561, p = .119]. Science-subject competition attendance [b_2 = .295, t (806) = 4.167, p < .001], respectively humanities competition attendance [b_3 = .421, t (806) = 7.041, p < .001] are significant predictors of academic achievement. Participation in sports [b_4 = -.094, t (806) = -.757, p = .450] and in art competitions [b_5 = .138, t (806) = 1.128, p = .260] are not significant predictors of academic results. Significant interaction effects of critical thinking and science subject competition attendance [b_6 = .065, t (806) = 3.115, p = .002], respectively critical thinking and humanities competition attendance [b_7 = .050, t (806) = 2.714, p = .007] on students' academic results were revealed. However, the interaction effects of critical thinking and sports competition attendance [b_8 = .029, t (806) = .948, p = .344], critical thinking and art competition attendance [b_9 = .037, t (806) = .898, p = .369] were not statistically significant. The moderating effect of type of competition on the relationship between critical thinking and academic mean results was statistically significant [F (4, 806) = 2.728, p = .028, ΔR^2= .012]. In the case of students who attended science subject competitions, with one unit increase of critical thinking, academic achievement also increases by .041 points [b = .041, t (806) = 2.932, p = .004], in the case of students who attended humanities competitions, one unit increase of critical thinking is associated with the improvement of academic results by .026 points [b = .026, t (806) = 2.608, p = .009] (Figure 2).

3.5.4.3. The moderating effect of conference attendance on the relationship between critical thinking and learning achievement

The overall moderation model including critical thinking ability as predictor variable, conference attendance as moderator variable and baccalaureate achievement as outcome variable, was statistically significant. Critical thinking and conference attendance accounted for a 9.3 % of the variance for baccalaureate results [F (3, 812) = 28.485, p < .001, R^2 = .093]. Students' critical thinking ability [b_1 = .077, t (812) = 8.607, p < 0.001] and participation in conferences [b_2 = .220, t (812) = 3.139, p =

0.001] are significant predictors of baccalaureate grades, but the moderating effect of conference attendance on the association between critical thinking and baccalaureate results is not statistically significant [F (1, 812) = 1.167, p = .280, ΔR^2 = .002].

The moderation model including critical thinking ability as predictor variable, conference attendance as moderator variable and academic achievement as outcome variable, was statistically significant. Critical thinking and conference attendance accounted for a 1.9 % of the variance for academic mean results [F (3, 812) = 6.402, p < .001, R^2 = .019]. Students critical thinking ability [b_1 = .025, t (812) = 3.648, p < 0.001] and participation in conferences [b_2 = .145, t (812) = 2.411, p = 0.016] are significant predictors of academic performance, but the moderating effect of conference attendance on the association between critical thinking and academic achievement is not statistically significant [F (1, 812) = .091, p = .763, ΔR^2 < .001].

3.5.4.4. The moderating effect of academic discipline on the relationship between critical thinking and learning achievement

The moderation model testing the moderating effect of academic discipline (multi-categorical variable with four levels: social sciences, humanities, natural sciences, sports) on the relationship between critical thinking and baccalaureate results is statistically significant [F (7, 808) = 15.650, p < .001, R^2 = .12].

Students' critical thinking predicted positively their baccalaureate results [b_1 = .056, t (808) = 4.253, p < 0.001], and sports academic discipline negatively predicted students' baccalaureate results [b_2 = -.607, t (808) = -3.904, p < 0.001]. The moderating effect of academic discipline on the association between critical thinking and baccalaureate results is not statistically significant [F (3, 808) = .532, p = .660, ΔR^2 = .002].

The model testing the moderating effect of academic discipline on the relationship between critical thinking and academic results is statistically significant [F (7, 808) = 12.266, p < .001, R^2 = .071].

Humanities [b_2 = .259, t (808) = 4.027, p < 0.001] and natural sciences [b_3 = .355, t (808) = 6.557, p < .001] academic disciplines significantly predicted students' academic achievement. The interaction effect of critical thinking and sports academic discipline was statistically significant [b_4 = .067, t (808) = 2.364, p = .018]. The moderating effect of academic discipline on the association between critical thinking and academic results is not statistically significant [F (3, 808) = 2.434, p = .064, ΔR^2 = .007].

3.5.4.5. The moderating effect of gender on the relationship between critical thinking and learning achievement

Students' critical thinking and gender accounted for a 9.3 % of the variance for baccalaureate results [F (3, 812) = 25.775, p < .001, R^2 = .093]. Students critical thinking ability [b_1 = .073, t (812) = 8.119, p < 0.001] and their gender [b_2 = -.174, t (812) = -2.325, p = 0.020] are significant predictors of baccalaureate performance, indicating that females have higher baccalaureate mean results than males, but the moderating effect of gender on the association between critical thinking and baccalaureate mean scores is not statistically significant [F (1, 812) = 3.315, p = .069, ΔR^2 < .004].

Students' critical thinking and gender accounted for a 2 % of the variance for academic performance [F (3, 812) = 5.837, p < .001, R^2 = .020]. Only students' critical thinking ability [b_1 = .022, t (812) = 3.220, p = 0.001] significantly predicted students' academic performance. The moderating effect of gender on the association between

critical thinking and academic mean scores is not statistically significant [F (1, 812) = 1.989, p = .159, $\Delta R^2 < .002$].

3.5.5. Discussion and Conclusions

The aim of the present study was the detailed investigation of some competition- and education-related moderator variables' effect on the relationship of one of the most important 21st century skills, namely undergraduate students' critical thinking and their learning achievement. This is the first study that measured the moderating effect of competition levels (school-level competitions, county- and national Olympiads), competition types (science subject, humanities, sports, art), conference attendance and academic disciplines (social science, humanities, natural science, sport) on the association between critical thinking and learning performance.

The model investigating the moderating effect of competition level on the association between critical thinking and baccalaureate mean results has shown that critical thinking, county Olympiad and national Olympiad attendance were significant predictors of baccalaureate achievement. The interaction effect of critical thinking and competition level was not statistically significant. The model testing the moderating effect of competition level on the relationship between critical thinking and academic mean results revealed that critical thinking, school-level competition, county Olympiad and national Olympiad attendance were significant predictors of academic achievement, respectively the interaction effect of critical thinking and competition level was statistically significant. Based on our results, the first hypothesis was partly supported, as competition levels were significant predictors of learning achievement, confirming the results of earlier studies that also found a positive relationship between competition, Olympiad attendance and learning achievement (Chen & Chang, 2020; Makhdum et al., 2023; Yang et al., 2020). Competition level was a significant moderator of the association between critical thinking and academic results.

The model investigating the moderating effect of competition type on the association between critical thinking and baccalaureate mean results demonstrated that critical thinking, science subject and humanities competition attendance were significant predictors of baccalaureate achievement, but the interaction effect of critical thinking and competition type was not statistically significant. Similarly, the model examining the moderating effect of competition type on the association between critical thinking and academic mean results revealed that science subject and humanities competition attendance were significant predictors of academic achievement, the interaction effect of critical thinking and competition type was statistically significant. The second hypothesis was partly supported by our results, because within competition types, science subject and humanities competition attendance were significant predictors of learning performance, and competition type is a significant moderator of the relationship between critical thinking and academic mean results. These results support the findings of earlier studies revealing the positive effect of science subject competitions or Olympiads on students' critical thinking (Akbar et al., 2022; Chiang et al., 2023; Greyling, 2023; Rif'at et al., 2022).

The models testing the moderating effect of conference attendance on the relationship between critical thinking and learning achievement revealed that critical thinking and conference attendance are significant predictors of baccalaureate and academic results, but the interaction effect of critical thinking and conference attendance was not statistically significant. These results highlight the importance of the organisation of extracurricular activities that could have a positive effect on students' learning achievement (Díaz-Iso et al., 2019; Feraco et al., 2022). We can

conclude that the first part of our third hypothesis is confirmed, because conference attendance is a significant predictor of learning achievement, but the second part of the hypothesis, regarding the moderating effect of conference attendance, is not supported by the results. That can be explained by a limitation of our study, namely that the majority of participants were 1st year undergraduate students and the majority of them did not participate in conferences.

Investigating the moderating effect of academic discipline on the association between critical thinking and baccalaureate mean results it was revealed that critical thinking and sports academic discipline (compared with social sciences profiles) were significant predictors of baccalaureate achievement, but the interaction effect of critical thinking and academic discipline was not statistically significant. Sports academic discipline compared with social science academic discipline negatively predicted students' baccalaureate grades, indicating that sport profile students have lower mean baccalaureate grades than social sciences students. Natural sciences and humanities academic disciplines (compared with social sciences profiles) were significant predictors of academic results, indicating that the natural science and humanities academic disciplines are associated with higher academic achievement (Ali & Awan, 2021; Ibrahim et al., 2021). A significant interaction effect of critical thinking and sports academic discipline was found, revealing that in the case of sports students the increase of critical thinking is associated with the increase of academic results. However, the overall moderating effect of academic discipline is not statistically significant, indicating that our fourth hypothesis is only partially supported, which could be explained by the unequal distribution of participants based on academic disciplines.

Gender is a significant predictor of baccalaureate results, females having higher achievement than males (Ali & Awan, 2021; Darmaji et al., 2022; Iqbal et al., 2021; Nwuba et al., 2022), but is not a significant moderator of the relationship between critical thinking and learning performance (Purba, 2022). The fifth hypothesis is partly supported by these results, which can be explained by the unequal gender distribution of students.

In order to reduce the limitations of this study, for the increase of its internal and external validity, several future directions and improvements are suggested like the replacement of the convenience sampling method with cluster sampling method, the extension of the sample to Romanian- and English-speaking students, the more detailed operationalisation of academic disciplines, the conduction of longitudinal studies, the measurement of students' critical thinking dispositions.

3.6. Study 5. The Improvement of Psychology and Special Education Students' Critical Thinking and Meta-Comprehension Accuracy through the Concept Mapping Method

3.6.1. Introduction

The improvement of higher order cognitive abilities, like critical thinking and metacognition, beside the traditional teaching methods, like the explanation of the material by the teacher and dictation, that require particularly memorisation, could be reached more effectively with methods that require the active participation of students in classes, the identification of the associations between the different items of information, the systematisation and synthesis of the learned material through deep information processing, like the concept mapping method (Barta et al., 2022; Carvalho et al., 2020; Mohammadi et al., 2019; Powell et al., 2021; Stevenson et al., 2017).

The detection of the relationships between the concepts contributes to the activation of metacognitive processes, to the improvement of meta-comprehension accuracy, as the hierarchical systematisation of the concepts promotes a more precise monitorisation of the acquired knowledge and comprehension, students evaluate their own knowledge more effectively, so they predict their academic or task performance more accurately (Powell et al., 2021; Prinz et al., 2020; Yang et al., 2022). The feedback on the actual performance, the discussion of the task also increases students' meta-comprehension accuracy (Callender et al., 2016; Carpenter et al., 2019; Stevenson et al., 2017). The high level of meta-comprehension accuracy leads to the increase of students' academic achievement, because the chance of over- or underestimating their own knowledge decreases, they detect more precisely those items of information that need to be deepened for the increase of their own achievement (Cai et al., 2019; Edossa et al., 2023; Hassan et al., 2022; Jansen et al., 2020; Muncer et al., 2021; Oppong et al., 2019; Souhila, 2022; Xue et al., 2021). The understanding of the associations between the different items of information activate students' critical thinking processes, like recognition of assumptions, induction, deduction, inference (Barta et al., 2022). Students identify more effectively the information, rules, premises needed for the resolving of given tasks and problems, that result in more precise and more objective conclusions (Barta et al., 2022; Carvalho et al., 2020; Khrais & Saleh, 2020; Silva et al., 2022; Roshangar et al., 2020).

The role of critical thinking and meta-comprehension accuracy is very important in interpreting and in estimating the validity of psychology-themed scientific works. The improvement of critical thinking abilities and meta-comprehension enables distinguishing between scientific and unscientific sources on the internet, drawing the correct conclusions from research results, acquiring information from multiple resources in the interest of the correct generalisation of the results (Beauvais, 2022). The aim of the present study beside the improvement of students' critical thinking and meta-comprehension accuracy, is the increase of the psychological scientific text comprehension ability with the concept mapping and feedback methods.

3.6.2. Aims of the Study and Hypotheses

The main aim of this study is testing the effect of the concept mapping method compared with the traditional learning method (rereading) for the improvement of Psychology and Special education students' critical thinking and meta-comprehension accuracy regarding scientific psychological texts. Our further aim is the investigation of the concept mapping method's effectiveness on students' psychological scientific text comprehension. Additionally, we also analyse the effect of giving feedback on concept maps on critical thinking, meta-comprehension accuracy and text comprehension achievement, as well as the differences between the rereading group (i.e. traditional learning, waiting list group) the feedback receiving group and the one receiving no feedback.

Based on the theoretical background the following hypotheses were formulated:

1. There is a significant improvement of critical thinking in the experimental groups (the concept mapping with and without feedback groups), whereas in the rereading group (waiting list group) there will be no significant improvement of critical thinking.
2. The concept mapping method has a significant effect on students' meta-comprehension accuracy; the use of the concept mapping method in the experimental groups (the concept mapping with and without feedback groups)

results in more accurate text comprehension judgments than in the rereading group (waiting list group).

3. The concept mapping method has a significant effect on students' text comprehension achievement; students in the concept mapping groups (with and without feedback) perform better at psychological text comprehension tests than students in the rereading group (waiting list group).

3.6.3. Methods

3.6.3.1. Participants

A total of 82 Psychology and Special Education students from Babeș-Bolyai University were randomly assigned to the concept mapping with feedback (n = 27), concept mapping without feedback (n = 28) or to the waiting list (rereading) group (n = 27). 23 participants (10 concept mapping with feedback, 9 concept mapping without feedback and 4 waiting list – rereading) did not complete post and follow-up measures of critical thinking. 59 participants (17 concept mapping with feedback, 19 concept mapping without feedback and 23 waiting list – rereading) were included in the data analysis regarding critical thinking (Figure 1), and 78 participants in data analysis regarding meta-comprehension. The attrition rate for critical thinking outcome in the concept mapping with feedback group was 37%, in the concept mapping without feedback group was 32.1% and in the waiting list (rereading group) was 14.8%, for meta-comprehension outcomes the attrition rate in the concept mapping with feedback and rereading groups was 7.4%, and in the concept mapping without feedback group was 0%. 93.6% of the participants were females, 89.7% of them studied Psychology. 61.5% were first year, 20.5% were second year and 17.9% were third year students, their mean age was 20.22 years (SD = 1.66).

3.6.3.2. Instruments

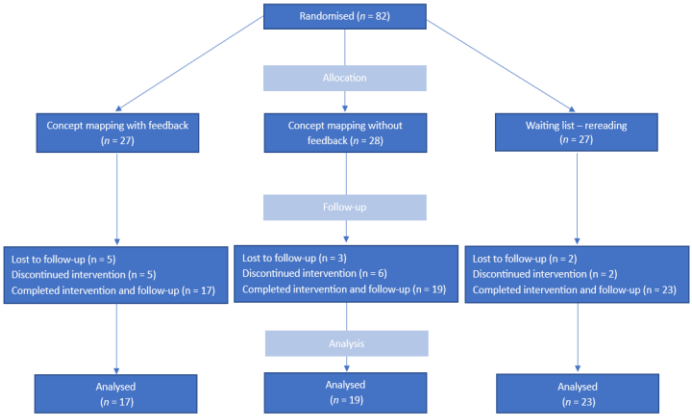
In a demographic questionnaire students' gender, age, academic discipline and study year were assessed.

For the measurement of students' critical thinking ability, a test was elaborated on the basis of two subscales, deduction and induction, of the Cornell Critical Thinking Test Level Z (Ennis et al., 2005), with the same instructions and the same internal structure. The deduction section consisted of 12, while the induction section of 10 items. The test has a forced-choice, dichotomous answer format.

The students created the concept maps based on five short texts, psychology research summaries with equal length and difficulty regarding the subject of psychological well-being during the Covid-19 pandemic. Their text comprehension achievement was assessed with five short tests regarding the read texts, with 5 questions and 4 possible answers. The following item was used for the measurement of meta-comprehension: ‚Estimate on a 6-point Likert scale your comprehension of the read text and your answering accuracy in the following test regarding this text: 0 - I didn't understand at all, I couldn't answer correctly any questions regarding this text, 6 - I understood it completely, I could answer correctly all the five questions regarding the text.

Figure 1

CONSORT Flowchart of Participants



Three indicators of meta-comprehension accuracy were computed. Absolute meta-comprehension accuracy was determined as the absolute error and as the confidence bias. Absolute error was calculated as the absolute difference between predicted and actual performance, and confidence bias as the signed difference between predicted and actual performance (Schraw, 2009; Wiley et al., 2016). Relative meta-comprehension accuracy was the third indicator. For each participant the intraindividual Pearson correlation between his or her prediction and actual achievement was calculated, higher correlation coefficients indicating better relative meta-comprehension accuracy across the texts (Griffin et al., 2008; Wiley et al., 2016; Zhang & Wang, 2014).

3.6.3.3. Research design

A randomised controlled trial, mixed 3 (method: concept mapping with feedback, concept mapping without feedback and waiting list – rereading) x 3 (measurement time: pretest, posttest and follow-up) factorial design was applied. The effect of independent variables (method and time) was tested on the following dependent variables: critical thinking, text comprehension achievement, prediction (judgement) regarding the text comprehension, meta-comprehension accuracy.

3.6.3.4. Procedure

Before the pretest students gave their informed consent to participate in the study. In the pretest the demographic questionnaire and the critical thinking test were completed. After the pretest the students were allocated randomly into one of the three groups: concept mapping with feedback, concept mapping without feedback and rereading (waiting list) group.

In the two concept mapping groups the first session was a two-hour training in the concept mapping method, practice in concept map construction: elaboration of a concept map by the students on the basis of an example text, the evaluation and discussion of an expert concept map elaborated based on the same text. During the next five sessions in the two experimental groups the students created a concept map

individually based on the given psychological text (in 30 minutes). After the elaboration of the concept maps, they answered the question regarding their meta-comprehension and after this, they completed the text comprehension test. In one of the experimental groups, the students received feedback at the end of the sessions in the form of interactive discussion of the text and the possible relationships among the concepts. In the rereading (waiting list) group students reread the text for 10 minutes, then they answered the question regarding their meta-comprehension and completed the text comprehension test. After the intervention period students completed the critical thinking test as a posttest, and two months after the posttest they also participated in a follow-up measurement of critical thinking. The rereading, waiting list group participated in the concept mapping with feedback method after the follow-up measurement.

3.6.3.5. Data analysis

A priori power analysis via G*Power3 (Faul et al., 2007) for mixed ANOVA based on type I error with a p-value of 0.05 and statistical power of 0.80 with three groups and three measurements showed that for a medium effect size ($f = 0.25$) the required sample size is n = 36; for one-way ANOVA based on type I error with a p-value of 0.05 and statistical power of 0.80 with three groups showed that for a medium effect size ($f = 0.25$) the required sample size is n = 159.

A mixed 3 (method: concept mapping with feedback, concept mapping without feedback, rereading) x 3 (measurement time: pretest, posttest, follow-up) ANOVA test was conducted to assess the effect of method and time on students' critical thinking. One-way ANOVA tests were conducted for the assessment of the effect of the method on students' text comprehension performance, predictive judgement and meta-comprehension accuracy.

3.6.4. Results

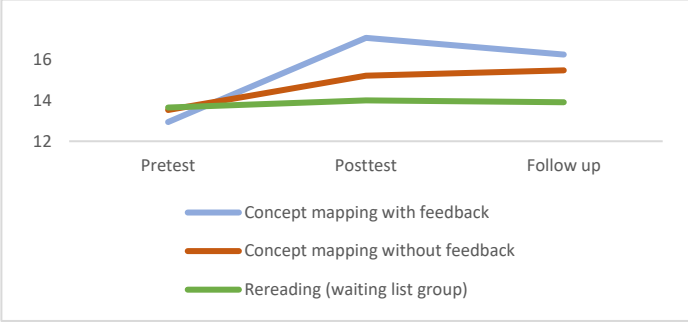
3.6.4.1. The effect of the concept mapping and feedback methods on students' critical thinking

A 3 x 3 mixed ANOVA test was conducted to assess the effect of time and the method on students' critical thinking. Preliminary analysis indicated that the assumptions of two-way mixed ANOVA were met.

The main effect of time was statistically significant, $F\,(2, 112) = 23.54$, $p < .001$, $\eta_p^2 = .296$, indicating that 29.6% of the variance for critical thinking was explained by the within-subjects variable, students having lower critical thinking scores in the pretest (M = 13.41, SD = 2.05), than in the post test (M = 15.27, SD = 2.41) and in follow-up measures (M = 15.08, SD = 2.25).

The main effect of the method was also statistically significant, $F\,(2, 56) = 5.52$, $p = .006$, $\eta_p^2 = .165$, indicating that 16.5% of the variance for critical thinking was explained by the between-subjects variable. Due to the unequal sample size of the groups Scheffe post hoc comparisons were conducted to evaluate pairwise differences between the levels of the between-subjects variable (method). Pairwise comparisons indicated a statistically significant difference in critical thinking between the concept mapping with feedback group (M = 15.41, SD = 1.82) and the waiting-list group (rereading group) (M = 13.85, SD = 2.08), p = .007. There was no statistically significant difference in critical thinking between the concept mapping with feedback and concept mapping without feedback (M = 14.74, SD = 2.23) groups, p = .401, neither between the concept mapping without feedback and the rereading groups, p = .169.

Figure 2
The effect of the concept mapping and feedback methods on students' critical thinking



Additionally, there was a statistically significant interaction between time and method, F (4, 112) = 6.462, $p < .001$, $\eta_p^2 = .188$, indicating that time and method accounted for a 18.8% of the variance for critical thinking. Students in the concept mapping with feedback group achieved higher scores of critical thinking at the posttest (M = 17.06, SD = 1.98) and at the follow-up (M = 16.24, SD = 1.60) than at the pretest (M = 12.94, SD = 1.89). Students in the concept mapping group without feedback also performed better on the critical thinking test at the posttest (M = 15.21, SD = 2.46) and at the follow-up (M = 15.47, SD = 1.95) measures than at the pretest (M = 13.53, SD = 2.27). Nevertheless, in the control (rereading) group the critical thinking scores were similar at the pretest (M = 13.65, SD = 2.01), the posttest (M = 14.00, SD = 1.81) and the follow-up (M = 13.91, SD = 2.41) measures (Figure 2).

3.6.4.2. The effect of the concept mapping and feedback methods on students' text comprehension performance, predictive judgements, and meta-comprehension accuracy

One-way ANOVA tests were conducted for the assessment of the effect of the concept mapping method and feedback on students' text comprehension performance, predictive judgements and meta-comprehension accuracy. Levene's test indicated that the assumption of homogeneity for text comprehension performance was met, F_{Levene} (2, 75) = .569, p = .569. The effect of the method was statistically significant for text comprehension achievement, F (2, 75) = 6.285, p = .003, η_p^2 = .144, indicating that 14.4% of the variance for text comprehension performance was explained by the applied text comprehension method. The results of the Scheffe post hoc test indicated a statistically significant difference in text comprehension scores between the concept mapping with feedback (M = 3.09, SD = .83) and rereading group (M = 2.42, SD = .70), p = .01. Similarly, there was a significant difference between the concept mapping without feedback (M = 3.05, SD = .71) and the rereading group, p = .013. No statistically significant difference was detected between the concept mapping (with and without feedback) groups, p = .983.

The assumption of homogeneity for the predictive judgement was met, F_{Levene} (2, 75) = .061, p = .941. The effect of the method was statistically significant for the predicted performance, F (2, 75) = 7.156, p = .001, η_p^2 = .16, indicating that 16% of the variance for the prediction regarding text comprehension was explained by the

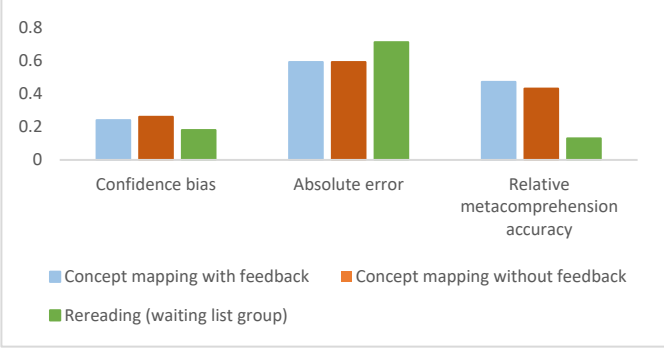
applied method. The results of the Scheffe post hoc test indicated a statistically significant difference in predictive judgements between the concept mapping with feedback (M = 3.32, SD = .79) and rereading groups (M = 2.61, SD = .77), p = .007. Similarly, there was a significant difference in the prediction scores between the concept mapping without feedback (M = 3.31, SD = .76) and the rereading group, p = .006. No statistically significant difference was detected between the concept mapping (with and without feedback) groups, p = .999.

The majority of students were overconfident, 61.5% of the confidence bias scores were higher than 0, 33.3% of the students underestimated their achievement on the text comprehension test.

The assumption of homogeneity for confidence bias, F_{Levene} (2, 75) = .647, p = .526, and for absolute error, F_{Levene} (2, 75) = 1.045, p = .357, was met. The effect of the method was not statistically significant for confidence bias, F (2, 75) = .064, p = .938, η_p^2 = .001, neither for absolute error, F (2, 75) = .509, p = .603, η_p^2 = .013. Based on the Levene's test result, the assumption of homogeneity for relative meta-comprehension accuracy was violated, F_{Levene} (2, 75) = 7.361, p = .001. The results of Welch's ANOVA were further analyzed, indicating that the applied method had a statistically significant effect on relative meta-comprehension accuracy, F_{Welch} (2, 44.715) = 4.704, p = .014.

Figure 3

The effect of the concept mapping and feedback methods on students' meta-comprehension accuracy



The Games-Howell post hoc test indicated a statistically significant difference in relative meta-comprehension accuracy between the concept mapping with feedback (M = .47, SD = .24) and rereading group (M = .13, SD = .51), p = .011. No statistically significant difference was detected between the concept mapping without feedback (M = .43, SD = .48) and the rereading groups, p = .077, neither between the concept mapping (with and without feedback) groups, p = .908 (Figure 3).

3.6.5. Discussion and Conclusions

The main aim of this study was the improvement of Psychology and Special education students' critical thinking ability, meta-comprehension accuracy and scientific text comprehension through the concept mapping and feedback methods.

Significant main effects of time and method on students' critical thinking ability were found, and a significant interaction effect of time and method was detected.

Students applying the concept mapping method and receiving group feedback performed significantly better on the posttest and the follow-up in the critical thinking test than the rereading group (Carvalho et al., 2020; Mohammadi et al., 2019; Roshangar et al., 2020; Tseng, 2019). The concept mapping group without feedback didn't achieve significantly better results in the critical thinking test than the rereading group, but there is a tendency of increase in their scores from the pretest to the posttest. In the rereading group, the critical thinking scores of students were similar each time they were measured. There was no statistically significant difference between the concept mapping with and without feedback groups in critical thinking. The first hypothesis is partially confirmed by the results, critical thinking in the concept mapping with feedback group showed significant improvement, while in the rereading group the results were similar during pretest, posttest and follow-up.

Similarly, a significant effect of the method on students' predictive judgements regarding their performance in the text comprehension test was observed. Students in the concept mapping groups predicted a significantly higher achievement than students in the rereading group. The concept mapping method increased their confidence regarding the comprehension of the texts. No statistically significant difference was detected between the two concept mapping groups in the predicted performance.

Our results showed no statistically significant difference in students' absolute meta-comprehension accuracy between the three groups, measured in absolute error and confidence bias. The majority of students were overconfident, 61.5% of the confidence bias scores were higher than 0, 33.3% of the students underestimated their achievement on the text comprehension test. However, a significant effect of the method was established on students' relative meta-comprehension accuracy. In the concept mapping group with feedback, the correlation between the predicted and actual performance, i.e. students' meta-comprehension accuracy was significantly higher than in the rereading group (Powell et al., 2021; Stevenson et al., 2017). No statistically significant difference was detected between the concept mapping without feedback and rereading groups, nor between the two concept mapping groups. The elaboration of the concept maps and the feedback, additional instructions, collective discussion of the read texts increased the effectiveness of the concept mapping method for the enhancement of students' meta-comprehension accuracy (Callender et al., 2016; Carpenter et al., 2019; Stevenson et al., 2017). Based on our results, the second hypothesis is partly accepted. No significant effect of the method was found on students' absolute meta-comprehension accuracy, but the concept mapping method with feedback had a significant positive effect on relative meta-comprehension accuracy.

Regarding students' text comprehension achievement, a significant effect of the concept mapping method was detected. Students in the two concept mapping groups (with and without feedback) performed significantly better at text comprehension tests than students in the rereading group. The concept mapping method is an effective learning method for the enhancement of students' comprehension (Powell et al., 2021), because it requires the activation of higher order cognitive abilities, like metacognitive strategies for the interpretation, deep analysis, and identification of relevant relationships in the read text (Stevenson et al., 2017). No statistically significant difference was found between the concept mapping with and without feedback groups in comprehension performance. Our third hypothesis is supported by the results, the text comprehension of the students in the concept mapping groups (with and without feedback) was better than that of the students in the rereading group.

The majority of the limitations of this study can be attributed to the convenience sampling method and replacing it with the stratified sampling method is suggested for future studies. The high attrition rate could also reduce the internal and external validity of the results. The improvement of students' intrinsic and extrinsic motivation for the increase of their active participation, and the optimisation of the time-effectiveness of the concept mapping method elaboration would also be meritorious in further studies. The online nature of the experiment, the absence of personal interaction could also explain the high attrition rate and the weaknesses of the applicability of the feedback method. A future direction is the application of the concept mapping and feedback methods in face-to-face format, and the comparison of the effectiveness of the individual, the collaborative, the paper-pencil and the digital concept mapping elaboration forms for the improvement of 21st century skills.

CHAPTER IV. GENERAL CONCLUSIONS AND IMPLICATIONS

4.1. General Conclusions

The **first main objective** of this thesis was to investigate the concept mapping method effectiveness for the improvement of students' critical thinking skills and dispositions, and to compare the effectiveness of this method with the traditional teaching method (Study 1). To address this objective, a meta-analysis was conducted, in which we also examined the impact of the concept mapping method on subcomponents of critical thinking skills and dispositions. Our overall results showed moderate effects of the concept mapping method compared with the control condition (traditional learning method) for critical thinking ability and disposition outcomes. A significant relationship between gender and critical thinking disposition ES was established, gender is a significant moderator of the effect of the concept mapping method on critical thinking disposition. Students' gender, age, educational level, the number of concept maps created by them, concept mapping elaboration methods (collaborative, digital) had no significant moderating effects on critical thinking ability ES. The allocation type was found to be a significant moderator for critical thinking ability ES. The allocation type was not a significant moderator for critical thinking disposition ES, which could be explained with the reduced number of studies measuring critical thinking disposition.

The **second main goal** of the thesis was the validation of the Metacognitive Awareness Inventory (MAI) (Harrison & Vallin, 2018; Moxon, 2022; Schraw, & Dennison, 1994) (**Study 2 A**) and the Cornell Critical Thinking Test Level Z (CCTT Level Z) (Alias et al., 2022; Ennis et al., 2005 Imperio et al., 2020 Leach et al., 2020) in Hungarian language (**Study 2 B**). Due to the lack of results regarding the factorial structure of the CCTT Level Z the main objective of Study 2 A was the comparison of several factorial models (correlated and hierarchical) translated into Hungarian language. The abbreviated 22-item four-factor second-order Hungarian version of the CCTT exceeded our criteria for excellent model fit (Hu, & Bentler, 1999; Kline, 2015). The results regarding the measurement invariance test of the 22-item Hungarian version of the CCTT indicated configural and scalar invariance across genders, the two tested models having acceptable fit indices for both the male and the female groups. Similarly, based on the contradictory results regarding the factorial structure of the MAI, the main objective of Study 2 B was the investigation and the comparison of the validity and factorial structure of the 52-item version of MAI proposed by Schraw and Dennison and the 19-item version of MAI proposed by Harrison and Vallin on a sample of Hungarian native language students. Based on the CFA results the 52-

item Schraw and Dennison model indicated unacceptable GFI, CFI and TLI values (Hu & Bentler, 1999), but the 19-item Harrison-Vallin model indicated acceptable and excellent model fit indices on every evaluated criterion (Awang, 2012; Hu & Bentler, 1999; Kline, 2015). Measurement invariance across genders was tested on our better fitted 19-item Harrison-Vallin model, applying a series of multigroup-CFAs. The four tested models (configural, metric, scalar, and residual) indicated acceptable model fit indices for both gender groups.

The third main aim of this doctoral thesis was to examine the differences in cognitive skills between students attending and not attending competitions, respectively the main and interaction effects of the participation of students in different types (science subject, humanities and social sciences, sports, art) and levels (school-level competition, county Olympiad and national Olympiad attendance) of competitions on their critical thinking and metacognition (**Study 3**). The results of the study indicated significant effects of the level and type of competitions on students' critical thinking skills and their metacognitive awareness. Significant differences were found in critical thinking between students not attending competitions and students attending Olympiads. Students attending science subject and humanities competitions had significantly higher critical thinking scores than students not attending competitions and students who participated in sports competitions. Students attending national level Olympiads had higher metacognitive knowledge than students who attended lower-level competitions.

The **fourth main goal** of the thesis was to test the predictive effect of critical thinking on students' learning achievement and the moderating effect of competition level, competition type, conference attendance, academic discipline, and gender on the association between critical thinking and learning performance (**Study 4**). The results of the moderation analyses regarding the moderating effect of competition level on the association between learning results (operationalised with baccalaureate and academic mean results) indicated that critical thinking, county Olympiad and national Olympiad attendance were significant predictors of baccalaureate and academic achievement. The interaction effect of critical thinking and competition level on baccalaureate results was not statistically significant, but the level of competition was a significant moderator of the association between critical thinking and academic results. The results of the moderation analyses regarding the moderating effect of competition type on the relationship between critical thinking and learning performance, revealed that critical thinking, science subject and humanities competition attendance were significant predictors of baccalaureate achievement and academic achievement, the interaction effect of critical thinking and competition type on academic achievement was statistically significant, but the interaction effect of critical thinking and competition type on baccalaureate performance was not statistically significant. The results of moderation analyses regarding the predictive and moderating effect of conference attendance on students' learning results, indicated that critical thinking and conference attendance are significant predictors of baccalaureate and academic results, but the interaction effect of critical thinking and conference attendance was not statistically significant. This insignificant result can be explained by the tendency that the majority of students do not participate in conferences in their first academic year. Examining the moderating effect of academic discipline on the association between critical thinking and baccalaureate mean results it was revealed that critical thinking and sports academic discipline (compared with social sciences profiles) were significant predictors of baccalaureate achievement, but the interaction effect of

critical thinking and academic discipline was not statistically significant. Natural sciences and humanities academic disciplines (compared with social sciences profiles) were significant predictors of academic results. A significant interaction effect of critical thinking and sports academic discipline was found. Nevertheless, the overall moderating effect of academic discipline is not statistically significant. The moderation analyses regarding the predictive and moderating effect of gender, revealed that gender is a significant predictor of baccalaureate results, females having higher achievement than males, but is not a significant moderator of the relationship between critical thinking and learning performance (Purba, 2022).

The **fifth main aim** of this thesis was to investigate the effect of the concept mapping method compared with the traditional learning method (rereading) for the improvement of Psychology and Special education students' critical thinking and meta-comprehension accuracy regarding scientific psychological texts (**Study 5**). The further aim of the Study 5 was the examination of the concept mapping method's effectiveness on students' psychological scientific text comprehension and to test the effect of giving feedback on concept maps on critical thinking, meta-comprehension accuracy and text comprehension performance. The results indicated that critical thinking in the concept mapping with feedback group showed significant improvement, while in the rereading group the results were similar during pretest, post-test and follow-up. No significant effect of the method was found on students' absolute meta-comprehension accuracy, but the concept mapping method with feedback had a significant positive effect on relative meta-comprehension accuracy. The text comprehension of the students in the concept mapping groups (with and without feedback) was better than that of the students in the rereading group.

4.2. Implications of the Study

4.2.1. Theoretical and Conceptual Implications

The present doctoral thesis approaches academic excellence as a characteristic of students with high learning achievement that is attainable through the accumulation of expertise with the active participation in supplementary activities, like participation in conferences, competitions, Olympiads, beside the compulsory education, which improve their 21st century cognitive abilities, like critical thinking and metacognition.

Study 1 is the first meta-analysis considering not only the effect of the concept mapping method on critical thinking ability and disposition, but also investigate the method's effectiveness on different subcomponents of critical thinking abilities and dispositions. The results highlight a moderate effect of the concept mapping method on students' critical thinking ability and disposition compared with the traditional learning method. The effect of concept mapping differs in diverse critical thinking ability and disposition subcomponents, it is more useful for the improvement of critical thinking abilities, such as recognition of assumptions, induction, deduction, inference, and dispositions, like analyticity, inquisitiveness, open-mindedness, truth-seeking, than the traditional teaching method. Similarly, this is the first meta-analysis that included studies conducted with students having different learning profiles and academic levels and treating these study-related variables as moderators. Previous meta-analysis restricted the included studies to profile-specific student samples (Romanko, 2016; Yue et al., 2017). Another innovation of this meta-analysis is the examination of the moderating effects of concept mapping elaboration methods (number of concept maps constructed, collaborative elaboration, computerised method), compared with previous meta-analyses that did not take into consideration the possible moderating effects of the characteristics of the concept mapping method

(Romanko, 2016; Yue et al., 2017). It is also a contribution to the literature that in our meta-analysis we included randomised, as well as non-randomised research, and the moderating effect of the randomisation process was also tested. Except for the randomisation process (for critical thinking ability) and gender (for critical thinking disposition), the tested moderating effects were not significant.

Study 3 is also an innovative one, being the first study that examined the effect of different types and levels of competitions on students' higher order cognitive abilities, respectively contributing to the literature with results regarding the effects of competition attendance on students' critical thinking and metacognitive awareness. The results of this study indicated the positive effects of competition attendance on students' 21st century cognitive skills, students participating in competitions having higher critical thinking ability and metacognitive knowledge than students not attending competitions. The critical thinking scores of students who participated in county or national Olympiads are higher than those of students who didn't participate in competitions. Students attending science subject, humanities and social sciences competitions achieved higher critical thinking scores than students not attending competitions and students attending sports competitions. Regarding the effect of competition level on students' metacognitive awareness, we found a significant main effect, but the effect of competition type on metacognition was not statistically significant.

The aim of Study 4 was the investigation of some competition- and education-related moderator variables' effect on the relationship of one of the most important 21st century skills, namely undergraduate students' critical thinking and their learning achievement. This is the first study that measured the moderating effect of competition levels (school-level competitions, county- and national Olympiads), competition types (science subject, humanities, sports, art), conference attendance and academic disciplines (social science, humanities, natural science, sport) on the association between critical thinking and learning performance. The results revealed the positive effect of participation in higher level competitions, in Olympiads, in conferences, on students' learning achievement. The association between critical thinking and academic achievement is stronger in the case of students who attended Olympiads from science subjects or humanities disciplines. Teachers, professors, educators, other academic employees, parents, and students could apply in practice the findings of this study that highlight the moderating and positive role of competitions and conferences on students' critical thinking and academic achievement.

Study 5 is the first study addressing the development of Psychology and Special education students' critical thinking, meta-comprehension accuracy and text comprehension achievement regarding psychological scientific texts, with the concept mapping and feedback methods. The results of the study highlighted that the critical thinking of students who elaborated concept maps and received feedback improved significantly, compared with the students who only reread the text. Providing feedback on the concept maps, the collective discussion of the read text improves the effectiveness of the concept mapping method for the enhancement of critical thinking. Students in the concept mapping with feedback and concept mapping without feedback groups demonstrated higher confidence regarding their text comprehension achievement. The three compared groups had similar results regarding confidence bias, the majority of students were overconfident regarding their text comprehension performance. The results also showed that the concept mapping method, together with the feedback method is effective for the enhancement of relative meta-comprehension

accuracy, calculated as the correlation between the predicted and actual text comprehension achievement of students. Students participating in the concept mapping groups, with or without feedback, also achieved better results at scientific text comprehension tests than students in the rereading group.

4.2.2. Methodological Implications

One of the main methodological implications of this thesis is the translation and validation into Hungarian language of the first instruments that measure students' critical thinking and metacognitive awareness. In Study 2 A our objective was to investigate the factorial structure of the CCTT Level Z. The CCTT Level Z is an extensively used instrument for the measurement of students' critical thinking in educational, cognitive psychology studies (Frost et al., 2019; Heidari, 2020; Kusumoto, 2018; Saud, 2020), however, there is a lack of empirical studies in the literature investigating the factorial structure of the CCTT Level Z. This is the first study that aimed at the comparison of different factorial structures of the CCTT level Z. Based on the recommendations of the authors of the test (Ennis et al., 2005), beside the unidimensional model, we also tested and compared several correlated and hierarchical factorial models (two-factor models: deductive reasoning, inductive reasoning; three-factor models: deduction, induction, meaning and fallacies; four-factor models: deduction, induction, meaning and fallacies, assumption identification) translated into Hungarian language. The results revealed that the four-factor structure of the test, including four higher order cognitive abilities within the general critical thinking factor, namely the deduction, meaning and fallacies, induction and assumption identification factors, is the most applicable, reliable and valid model for the measurement of critical thinking of undergraduate Hungarian-speaking students from Babeș-Bolyai University. Post hoc inspection of the four-factor second-order structure indicated a 22-item shortened version of the test with excellent fit indices. The results of the invariance test revealed configural and scalar invariance across genders of the 22-item four-factor structure Hungarian version of the CCTT. The translated and validated Hungarian version of the CCTT level Z, is a very important instrument for the empirical measurement of critical thinking skills in education, and in different work environments. The empirical measurement of students' critical thinking provides information about their level of thinking skills and promotes the identification of those skills that require further development in educational context. Study 2 B is the first study that compared the validity and factorial structure of the 52-item version of MAI proposed by Schraw and Dennison (1994) and the 19-item version of MAI proposed by Harrison and Vallin (2018), because of the contradictory results regarding the MAI's factor and item structure. The results demonstrated that the 19-item two-factor structure (metacognitive knowledge and metacognitive regulation) of MAI developed by Harrison and Vallin, compared with the 52-item version proposed by Schraw and Dennison, proved to be more valid and reliable, having better model fit indices on Hungarian native language undergraduate student sample, and it is reliably applicable for the measurement of students' metacognitive knowledge and regulation, independently of gender.

Study 5 brings methodological contributions to the literature as well. Firstly, the summaries of the psychological empirical studies regarding the topic of psychological well-being during the COVID-19 pandemic, read by the students, can be applied in future studies or in educational practice as a part of the assessment of meta-comprehension accuracy, or text comprehension, or in different Psychology courses, or seminars. Secondly, the elaborated text comprehension tests, can also be used in

future research or in educational context as an objective instrument for the measurement of students' scientific psychological text comprehension achievement. Finally, we also adapted two subscales (deduction and induction) of Cornell Critical Thinking Test Level Z for psychological topics, retaining its internal structure and instructions. The critical thinking test regarding psychological research and concepts can be applied in future studies for the measurement of students' deduction and induction.

4.2.3. Practical Implications

The findings of the meta-analysis in Study 1 draw attention to the effectiveness of the concept mapping method as an active learning, metacognitive regulation strategy for teaching and applying it in educational settings. The deep processing of course material and the exploration of the relationships between relevant concepts presuppose the functioning of metacognitive processes, which are activated using the concept mapping method (Khine et al., 2019, Powell et al., 2021; Stevenson et al., 2017). The method can be applied in a variety of ways, either in paper-pencil or digital format, individually or in groups, and the introduction of the method to students is time and energy efficient. Providing feedback on the created concept maps further enhances the effectiveness of the method in promoting cognitive processes (Kaddoura et al., 2016; Powell et al., 2021; Stevenson et al., 2017), as well as academic performance (Dmoshinskaia et al., 2021; Joseph et al., 2017).

The findings of Study 5 are very important from a practical point of view as well. The application of the learned concept mapping method during the experiment resulted in the development of students' critical thinking, respectively in higher text comprehension achievement and meta-comprehension accuracy. Similarly, the use of feedback in the form of group discussion regarding the concept maps and the interpretation of the texts resulted in higher levels of critical thinking, text comprehension performance and meta-comprehension accuracy. The results of this study are usable in educational, teaching context, highlighting that the concept mapping method, giving feedback on the concept maps, the interactive discussion of the read scientific information, the clarification of the associations between the scientific concepts develop students' critical thinking, have a positive effect on their meta-comprehension accuracy, on the estimation of the profoundness of the information processing and of their own performance, respectively on their text comprehension achievement.

4.3. Limitations and Future Directions

Beside the theoretical, methodological, and practical implications of our results, the present thesis has some general limitations that should be considered in the interpretation and generalisation of the results and may serve as a guide for future research. First, the number of studies included in the meta-analysis regarding the concept mapping method effectiveness for students' critical thinking disposition is reduced, which can affect the generalisability of the results. In the moderator analysis some subgroups were underrepresented due to the reduced number of studies involving secondary school students, as well as digital and collaborative concept mapping elaboration methods. Based on these limitations, future studies should focus on the examination of the concept mapping method's effectiveness on students' critical thinking dispositions, not only in the case of college students, but also in secondary education. It would be meritorious to investigate the method's effectiveness, considering the expertise, the earlier experience in the application of the method. Future studies could test the method's effectiveness in the case of students with

different science and humanities disciplines and could apply the digital and collaborative forms of the method for improving students' critical thinking. The comparison of the traditional, paper-pencil concept mapping method with the digital and collaborative forms, and with other active learning methods such as mind mapping, argument mapping, flipped learning model, problem-based learning model would also be beneficial.

A general limitation of the thesis is the convenience sampling method that could have effects on the generalisability of the results of each study. In Studies 2 A and 2 B the instruments were applied, and their factorial structure was tested only on Hungarian-speaking undergraduate students from Babeș-Bolyai University which reduces the generalisability of the results. Future studies should involve students studying at other universities, master students and students with different academic profiles in the interest of confirming our results regarding the validity and factorial structure of the instruments. The results of the invariance tests should be interpreted carefully due to the unequal gender distribution of the sample. Besides the unequal distribution based on students' demographic characteristics (Studies 2 and 5), the excellence- and competition-related distribution in Studies 3 and 4 is also unequal. Future studies should apply the cluster sampling method for the equal and balanced representation of students based on demographic and excellence-related variables.

The self-reported format of the MAI is also a limitation of the thesis that may result in biases like social desirability or acquiescence. A possible explanation for the statistically insignificant results in Study 3 regarding the effects of competition level on students' metacognitive regulation, and of competition type on metacognitive knowledge and regulation, could be the self-reported format of the MAI, the answers and judgements of the participants regarding their metacognitive knowledge and applied metacognitive strategies could affect the results by the possible presence of social desirability bias or imprecise conclusions. Future studies should apply on-line measures of metacognition, beside the off-line measurement of metacognition with the MAI, giving the opportunity to test the convergent validity of the MAI.

The type of research designs of Studies 3 and 4 have limitations as well. In Study 3 a lack of experimental control is present; the level of students' metacognition and critical thinking were not examined before they participated in competitions. The ex post facto design also enhances the presence of latent variables that could explain the differences between students in the measured cognitive abilities. It would be worthwhile to conduct quasi-experiments, more specifically time series experiments for the analysis of the effect of different types and levels of competitions and conference participation on students' cognitive abilities with multiple pretest and post-test measures. Due to the cross-sectional, correlational design of Study 4 we can draw conclusions only regarding the linear relationships between the measured variables due to the lack of experimental control among the predictor and moderator variables. Secondly, students' critical thinking, their academic achievement, respectively conference participation could change in time. Students' critical thinking was measured minimum one year after the baccalaureate exam, their critical thinking ability may have improved during their academic studies. Similarly, with the increase of the expertise in academic learning their academic achievement could also increase. In the future it would be meritorious to conduct longitudinal studies for the multiple measurement of students' critical thinking, academic results, and conference participation. The application of multiple moderation analysis, the elaboration of a theoretical model regarding the effect of the tested moderator variables on the

relationship between critical thinking and learning achievement, is also a future direction.

A further limitation of Study 5 is the small sample size and high attrition rate in the concept mapping groups. The repetition of the study with a larger sample size is recommended. In the interest of lower attrition rates, it is recommended to include more incentives as well as to optimise the sessions to be more time-effective and more interactive, more interesting for the students. The face-to-face, personal meetings, sessions with the participants for the enhancement of communication and discussion could increase the effectiveness of the feedback.

Beside the above-mentioned future directions of this thesis, our further aim is the examination of students' critical thinking disposition, and to investigate the effect of the concept mapping method on students' critical thinking disposition. Besides the cognitive processes that influence students' academic excellence, the investigation of motivational (achievement motivation, intrinsic motivation) and emotional characteristics (achievement anxiety, emotion regulation) that may have an effect on their academic success is also a future research opportunity. The development of a logical thinking test for the measurement of students' metacognitive planning ability would also be meritorious.

BIBLIOGRAPHY

Abdelrahman, R. M. (2020). Metacognitive awareness and academic motivation and their impact on academic achievement of Ajman University students. *Heliyon*, *6*(9), e04192. https://doi.org/10.1016/j.heliyon.2020.e04192

Abu Bakar, M. A., & Ismail, N. (2020). EXPLORING STUDENTS' METACOGNITIVE REGULATION SKILLS AND MATHEMATICS ACHIEVEMENT IN IMPLEMENTATION OF 21ST CENTURY LEARNING IN MALAYSIA. *Problems of Education in the 21st Century*, *78*(3), 314–327. https://doi.org/10.33225/pec/20.78.314

Abueita, J. D., Al Fayez, M. Q., Alsabeelah, A., & Humaidat, M. A. (2022). The Impact of (STEAM) Approach on the Innovative Thinking and Academic Achievement of the Educational Robot Subject among Eighth Grade Students in Jordan. *Journal of Educational and Social Research*, *12*(1), 188-203. https://doi.org/10.36941/jesr-2022-0016

Akin, A., Abaci, R., & Cetin, B. (2007). The Validity and Reliability of the Turkish Version of the Metacognitive Awareness Inventory. *Educational Sciences: Theory & Practice, 7*(2), 671-678.

Akpur, U. (2020). Critical, Reflective, Creative Thinking and Their Reflections on Academic Achievement. *Thinking Skills and Creativity*, *37*, 100683. https://doi.org/10.1016/j.tsc.2020.100683

Alfayoumi, I. (2018). The impact of combining concept-based learning and concept mapping pedagogies on nursing STUDENTS' clinical reasoning abilities. *Nurse Education Today, 72,* 40-46. https://doi.org/10.1016/j.nedt.2018.10.009

Ali, G., & Awan, R.-N. (2021). Thinking based Instructional Practices and Academic Achievement of Undergraduate Science Students: Exploring the Role of Critical Thinking Skills and Dispositions. *Journal of Innovative Sciences*, *7*(1). https://doi.org/10.17582/journal.jis/2021/7.1.56.70

Alias, A., Mohtar, L. E., Ayop, S. K., & Rahim, F. R. (2022). A Systematic Review on Instruments to Assess Critical Thinking & Problem-Solving Skills. *EDUCATUM Journal of Science, Mathematics and Technology*, *9*, 38-47. https://doi.org/10.37134/ejsmt.vol9.sp.5.2022

Anwarudin, M., Dafik, & Ridlo, Z. R. (2021). The analysis of olympiad student's metacognition skills in solving the national sciences olympiad problem on two-variables linear equation system material. *Journal of Physics: Conference Series, 1832*(1), 012042. https://doi.org/10.1088/1742-6596/1832/1/012042

Appaw, E. L., Owusu, E., & Frimpong, R. (2021). Effect of Concept Mapping on the Achievement of Biology Students at the Senior High School Level in Ghana. *European Journal of Research and Reflection in Educational Sciences, 9*(2), 15-28.

Arbuckle, J. L. (2016). *IBM® SPSS® Amos™ 24 User's Guide*. IBM Corp.

Awang, Z. (2012). *Research methodology and data analysis.* Penerbit Universiti Teknologi MARA Press.

Barta, A., Fodor, L. A., Tamas, B., Szamoskozi, I. (2022). The development of students critical thinking abilities and dispositions through the concept mapping learning method – A meta-analysis. *Educational Research Review, 37*, 100481. https://doi.org/10.1016/j.edurev.2022.100481

Bayındır, D., Acar, İ. H., Yavuz, E. A., & Ahmetoğlu, E. (2021). To Compete or Not Compete: Contributions of Children's Regulation and Gender to Their Competitive Behaviors. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-021-01194-1

Beauvais, C. (2022). Fake news: Why do we believe it? *Joint Bone Spine*, *89*(4), 105371. https://doi.org/10.1016/j.jbspin.2022.105371

Boran, M., & Karakuş, F. (2022). The Mediator Role of Critical Thinking Disposition in the Relationship between Perceived Problem-Solving Skills and Metacognitive Awareness of Gifted and Talented Students. *Participatory Educational Research*, *9*(1), 61–72. https://doi.org/10.17275/per.22.4.9.1

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Bransen, D., Govaerts, M. J. B., Panadero, E., Sluijsmans, D. M. A., & Driessen, E. W. (2022). Putting self-regulated learning in context: Integrating self-, co-, and socially shared regulation of learning. *Medical Education*, *56*(1), 29–36. https://doi.org/10.1111/medu.14566

Butler, H. A., Pentoney, C., & Bong, M. P. (2017). Predicting real-world outcomes: Critical thinking ability is a better predictor of life decisions than intelligence. *Thinking Skills and Creativity*, *25*, 38–46. https://doi.org/10.1016/j.tsc.2017.06.005

Cai, Y., King, R. B., Law, W., & McInerney, D. M. (2019). Which comes first? Modeling the relationships among future goals, metacognitive strategies and academic achievement using multilevel cross-lagged SEM. *Learning and Individual Differences*, *74*, 101750. https://doi.org/10.1016/j.lindif.2019.06.004

Callan, G., Yang, N.-J., Zhang, Y., & Sciuchetti, M. B. (2020). Narrowing the Research to Practice Gap: A Primer to Self-Regulated Learning Application

in School Psychology. *Contemporary School Psychology*. https://doi.org/10.1007/s40688-020-00323-8

Callender, A. A., Franco-Watkins, A. M., & Roberts, A. S. (2016). Improving metacognition in the classroom through instruction, training, and feedback. *Metacognition and Learning*, *11*(2), 215–235. https://doi.org/10.1007/s11409-015-9142-6

Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-General Enhancements of Metacognitive Ability Through Adaptive Training. *Journal of Experimental Psychology: General, 148*(1), 51-64. http://dx.doi.org/10.1037/xge0000505

Carvalho, D. P. de S. R. P., Vitor, A. F., Cogo, A. L. P., Bittencourt, G. K. G. D., Santos, V. E. P., & Ferreira Júnior, M. A. (2020). Critical thinking in nursing students from two Brazilian regions. *Revista Brasileira de Enfermagem*, *73*(1), e20170742. https://doi.org/10.1590/0034-7167-2017-0742

Chen, C., Yang, C., Huang, K., & Yao, K. (2020). Augmented reality and competition in robotics education: Effects on 21st century competencies, group collaboration and learning motivation. *Journal of Computer Assisted Learning*, *36*(6), 1052–1062. https://doi.org/10.1111/jcal.12469

Chen, F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504. https://doi.org/10.1080/10705510701301834

Chen, M. A., & Hwang, G. (2020). Effects of a concept mapping-based flipped learning approach on EFL students' English speaking performance, critical thinking awareness and speaking anxiety. *British Journal of Educational Technology*, *51*(3), 817–834. https://doi.org/10.1111/bjet.12887

Chen, S. Y., & Chang, Y.-M. (2020). The impacts of real competition and virtual competition in digital game-based learning. *Computers in Human Behavior*, *104*, 106171. https://doi.org/10.1016/j.chb.2019.106171

Chiang, F.-K., Zhang, Y., & Lu, Y. (2023). Development and validation of a questionnaire for assessing perspectives of World Robot Olympiad on participants. *Research and Practice in Technology Enhanced Learning*, *18*, 016. https://doi.org/10.58459/rptel.2023.18016

Chiou, C.-C., Tien, L.-C., & Tang, Y.-C. (2020). Applying structured computer-assisted collaborative concept mapping to flipped classroom for hospitality accounting. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *26*, 100243. https://doi.org/10.1016/j.jhlste.2020.100243

Chu, S. K. W., Reynolds, R. B., Tavares, N. J., Notari, M., & Lee, C. W. Y. (2017). *21st Century Skills Development Through Inquiry-Based Learning*. Springer Singapore. https://doi.org/10.1007/978-981-10-2481-8

Cogliano, M., Bernacki, M. L., & Kardash, C. M. (2021). A metacognitive retrieval practice intervention to improve undergraduates' monitoring and control processes and use of performance feedback for classroom learning. *Journal of Educational Psychology*, *113*(7), 1421–1440. https://doi.org/10.1037/edu0000624

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112*(1), 155-159. http://dx.doi.org/10.1037/0033-2909.112.1.155

Concina, E. (2019). The Role of Metacognitive Skills in Music Learning and Performing: Theoretical Features and Educational Implications. *Frontiers in Psychology*, *10*, 1583. https://doi.org/10.3389/fpsyg.2019.01583

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*, 1-9.

Craig, K., Hale, D., Grainger, C., & Stewart, M. E. (2020). Evaluating metacognitive self-reports: Systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning*, *15*(2), 155–213. https://doi.org/10.1007/s11409-020-09222-y

Cui, L., Zhu, Y., Qu, J., Tie, L., Wang, Z., & Qu, B. (2021). Psychometric properties of the critical thinking disposition assessment test amongst medical students in China: A cross-sectional study. *BMC Medical Education*, *21*(1), 10, 1-8. https://doi.org/10.1186/s12909-020-02437-2

Daley, B. J., & Torre, D. M. (2010). Concept maps in medical education: An analytical literature review. *Medical Education*, *44*(5), 440–448. https://doi.org/10.1111/j.1365-2923.2010.03628.x

Darmaji, D., Kurniawan, D. A., Astalini, A., & Setiya Rini, E. F. (2022). Science Processing Skill and Critical Thinking: Reviewed Based on the Gender. *JPI (Jurnal Pendidikan Indonesia)*, *11*(1), 133–141. https://doi.org/10.23887/jpi-undiksha.v11i1.35116

Davies, M., & Barnett, R. (Eds.). (2015). *The Palgrave Handbook of Critical Thinking in Higher Education.* Palgrave Macmillan US. https://doi.org/10.1057/9781137378057

de Boer, H., Donker, A. S., Kostons, D. D. N. M., & van der Werf, G. P. C. (2018). Long-term effects of metacognitive strategy instruction on student academic performance: A meta-analysis. *Educational Research Review*, *24*, 98–115. https://doi.org/10.1016/j.edurev.2018.03.002

Deliligka, A., & Calfoglou, C. (2022). Using poetry to foster Critical Thinking and Metacognition in a Primary School EFL context. *Research Papers in Language Teaching and Learning, 12*(1), 167-187.

Díaz-Iso, A., Eizaguirre, A., & García-Olalla, A. (2019). Extracurricular Activities in Higher Education and the Promotion of Reflective Learning for Sustainability. *Sustainability*, *11*(17), 4521. https://doi.org/10.3390/su11174521

Dindar, M., Järvelä, S., & Järvenoja, H. (2020). Interplay of metacognitive experiences and performance in collaborative problem solving. *Computers & Education*, *154*, 103922. https://doi.org/10.1016/j.compedu.2020.103922

Dmoshinskaia, N., Gijlers, H., & de Jong, T. (2021). Learning from reviewing peers' concept maps in an inquiry context: Commenting or grading, which is better? *Studies in Educational Evaluation*, *68*, 100959. https://doi.org/10.1016/j.stueduc.2020.100959

Doleck, T., Bazelais, P., Lemay, D. J., Saxena, A., & Basnet, R. B. (2017). Algorithmic thinking, cooperativity, creativity, critical thinking, and problem solving: Exploring the relationship between computational thinking skills and academic performance. *Journal of Computers in Education*, *4*(4), 355–369. https://doi.org/10.1007/s40692-017-0090-9

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Edossa, A. K., Lockl, K., & Weinert, S. (2023). Developmental Relationship Between Metacognitive Monitoring and Reading Comprehension. *Journal of Educational and Developmental Psychology*, *13*(1), 1. https://doi.org/10.5539/jedp.v13n1p1

Efklides, A., & Misailidi, P. (2019). Emotional Self-Regulation in the Early Years: The Role of Cognition, Metacognition and Social Interaction. In D. Whitebread, V. Grau, K. Kumpulainen, M. M. McClelland, N. E. Perry, & D. Pino-Pasternak, *The SAGE Handbook of Developmental Psychology and Early Childhood Education* (pp. 502–514). SAGE Publications Ltd. https://doi.org/10.4135/9781526470393.n29

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 639-634. https://doi.org/10.1136/bmj.315.7109.629

Ennis, R. H. (2018). Critical Thinking Across the Curriculum: A Vision. *Topoi*, *37*(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4

Ennis, R. H., Millman, J., & Tomko, T. N. (2005). *Cornell Critical Thinking Tests Level X and Z Manual* (5th ed.). The Critical Thinking Co.

Ersteniuk, Y., Gasyuk, I., Boryschak, A., & Yakubovskyi, P. (2020). Methodology of Problems Creation and Selection for Astronomy Olympiads on Example of Tasks on the Topic of Kepler's Laws. *Journal of Vasyl Stefanyk Precarpathian National University*, *7*(1), 156–165. https://doi.org/10.15330/jpnu.7.1.156-165

Eshmirzaeva, M. A. (2020). Personality-oriented Approach to Education and Modern Pedagogical Technologies. *European Journal of Research and Reflection in Educational Sciences, 8*(8), 25-30.

Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction—The Delphi report*. California Academic Press.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). GPower 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/BF03193146

Feraco, T., Resnati, D., Fregonese, D., Spoto, A., & Meneghetti, C. (2022). An integrated model of school students' academic achievement and life satisfaction. Linking soft skills, extracurricular activities, self-regulated learning, motivation, and emotions. *European Journal of Psychology of Education*, *38*(1), 109–130. https://doi.org/10.1007/s10212-022-00601-4

Flavell, J. H. (1987). Speculation about the nature and development of metacognition. In F. Weinert, & R. Kluwe (Eds.). *Metacognition, motivation, and understanding* (pp. 21 - 29). Lawrence Erlbaum.

Fujiyama, H., Kamo, Y., & Schafer, M. (2021). Peer effects of friend and extracurricular activity networks on students' academic performance. *Social Science Research*, *97*, 102560. https://doi.org/10.1016/j.ssresearch.2021.102560

Garwood, J. K., Ahmed, A. H., & McComb, S. A. (2018). The Effect of Concept Maps on Undergraduate Nursing Students' Critical Thinking. *Nursing Education Perspectives*, *39*(4), 208–214. https://doi.org/10.1097/01.NEP.0000000000000307

Genç, G. (2017). The Relationship Between Academic Achievement, Reading Habits And Critical Thinking Dispositions of Turkish Terttiary Level EFL Learners. *Educational Research Quarterly, 41*(2), 43-73.

Greyling, J. (2023). Coding Unplugged—A Guide to Introducing Coding and Robotics to South African Schools. In J. Halberstadt, A. Alcorta de Bronstein, J. Greyling, & S. Bissett (Eds.), *Transforming Entrepreneurship Education* (pp. 155–174). Springer International Publishing. https://doi.org/10.1007/978-3-031-11578-3_9

Griffin, T. D., Jee, B. D., & Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory & Cognition*, *37*(7), 1001–1013. https://doi.org/10.3758/MC.37.7.1001

Gunawardena, M., & Wilson, K. (2021). Scaffolding students' critical thinking: A process not an end game. *Thinking Skills and Creativity*, *41*, 100848. https://doi.org/10.1016/j.tsc.2021.100848

Haber, J. (2020). *Critical Thinking*. MIT Press

Halpern, D. F. (1997). *Critical thinking across the curriculum: A brief edition of thought and knowledge*. Lawrence Erlbaum Associates.

Halpern, D. F., & Sternberg, R. J. (2020). An Introduction to Critical Thinking: Maybe It Will Change Your Life. In R. J. Sternberg & D. F. Halpern (Eds.), *Critical Thinking in Psychology* (2nd ed., pp. 1–9). Cambridge University Press. https://doi.org/10.1017/9781108684354.002

Harrison, G. M., & Vallin, L. M. (2018). Evaluating the Metacognitive Awareness Inventory using empirical factor-structure evidence. *Metacognition and Learning, 13*(1), 15–38. https://doi.org/10.1007/s11409-017-9176-z

Hashmi, A., Khalid, M., & Shoaib, A. (2019). A Cross-Sectional Study of Assessing Metacognitive Knowledge and Metacognitive Regulatory Skills among Prospective Teachers and Its Relation to their Academic Achievement. *Bulletin of Education and Research, 41*(2), 215-234.

Hassan, S., Venkateswaran, S. P., Agarwal, P., Sulaiman, A. R. B., & Burud, I. A. S. (2022). *Metacognitive Awareness And Its Relation To Students' Academic Achievement: Time To Ponder Its Implication In Delivery of The Curriculum.* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-1266966/v1

Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis. A Regression-Based Approach* (3rd ed.). Guilford Press.

Heidari, K. (2020). Critical thinking and EFL learners' performance on textually-explicit, textually-implicit, and script-based reading items. *Thinking Skills and Creativity*, *37*, 100703. https://doi.org/10.1016/j.tsc.2020.100703

Higgins, J. P. T. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Ho, V., Kumar, R. K., & Velan, G. (2014). Online testable concept maps: benefits for learning about the pathogenesis of disease. *Medical Education, 48,* 687-697. https://doi.org/10.1111/medu.12422

Hong, J.-C., Chen, M.-Y., & Hwang, M.-Y. (2013). Vitalizing creative learning in science and technology through an extracurricular club: A perspective based on activity theory. *Thinking Skills and Creativity*, *8*, 45–55. https://doi.org/10.1016/j.tsc.2012.06.001

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, L.-Y., & Yeh, Y.-C. (2017). Meaningful Gamification for Journalism Students to Enhance Their Critical Thinking Skills: *International Journal of Game-Based Learning*, *7*(2), 47–62. https://doi.org/10.4018/IJGBL.2017040104

Hwang, G.-J., Zou, D., & Lin, J. (2020). Effects of a multi-level concept mapping-based question-posing approach on students' ubiquitous learning performance and perceptions. *Computers & Education*, *149*, 103815. https://doi.org/10.1016/j.compedu.2020.103815

Ibrahim, I., Akmal, N., & Marwan, M. (2021). Problem-based Learning and Student Critical Thinking to Improve Learning Achievement at Private Universities in Aceh. *AL-ISHLAH: Jurnal Pendidikan*, *13*(2), 1142–1151. https://doi.org/10.35445/alishlah.v13i2.698

Idiege, K. J., & Nja, C. O. (2021). INTERACTIVE ACTIVATION AND COMPETITION (IAC) MODEL AS A METACOGNITIVE STRATEGY AND STUDENTS' ACADEMIC ACHIEVEMENT IN ELECTROCHEMISTRY IN CROSS RIVER STATE, NIGERIA. *Inter-Disciplinary Journal of Science Education (IJ-SED), 3*(1), 95-105.

Imperio, A., Kleine Staarman, J., & Basso, D. (2020). Relevance of the socio-cultural perspective in the discussion about critical thinking. *Journal of Theories and Research in Education, 15*, 1-19. https://doi.org/10.6092/ISSN.1970-2221/9882

Iqbal, M. Z., Khan, M. J., Javed, T., Rao, U., & Shams, J. A. (2021). Relationship Between Secondary Students' Critical Thinking and Academic Achievement: A Case of Public Schools. *Humanities & Social Sciences Reviews*, *9*(3), 1166–1174. https://doi.org/10.18510/hssr.2021.93115

Işiklar, S., & Abali-Öztürk, Y. (2022). The effect of philosophy for children (P4C) curriculum on critical thinking through philosophical inquiry and problem solving skills. *International Journal of Contemporary Educational Research*. *9*(1), 130-142. https://doi.org/10.33200/ijcer.942575

Jansen, R. S., van Leeuwen, A., Janssen, J., Conijn, R., & Kester, L. (2020). Supporting learners' self-regulated learning in Massive Open Online Courses. *Computers & Education*, *146*, 103771. https://doi.org/10.1016/j.compedu.2019.103771

Jatmiko, D. D. H., Fatahillah, A., Oktavianingtyas, E., & Rochmah, O. A. (2020). An analysis of olympiad students' critical thinking in solving National Science Olympiad (OSN) problem on number theory material. *Journal of Physics: Conference Series*, *1563*(1), 012037. https://doi.org/10.1088/1742-6596/1563/1/012037

Jordalen, G., Lemyre, P.-N., & Durand-Bush, N. (2019). Interplay of motivation and self-regulation throughout the development of elite athletes. *Qualitative Research in Sport, Exercise and Health*, *12*(3), 377–391. https://doi.org/10.1080/2159676X.2019.1585388

Joseph, C., Conradsson, D., Nilsson Wikmar, L., & Rowe, M. (2017). Structured feedback on students' concept maps: The proverbial path to learning? *BMC Medical Education, 17*(1), 90. https://doi.org/10.1186/s12909-017-0930-3

Karatas, K., & Arpaci, I. (2021). The Role of Self-directed Learning, Metacognition, and 21st Century Skills Predicting the Readiness for Online Learning. *Contemporary Educational Technology*, *13*(3), ep300. https://doi.org/10.30935/cedtech/10786

Khrais, H., & Saleh, A. M. (2020). The Effect of Concept Mapping on Critical Thinking of Jordanian Nursing Students. *Creative Nursing*, *26*(1), e19–e24. https://doi.org/10.1891/1078-4535.26.1.e19

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kravchenko, Z., & Nygård, O. (2022). Extracurricular activities and educational ouctomes: Evidence from high-performing schools in St Petersburg, Russia. *International Studies in Sociology of Education*, 1–20. https://doi.org/10.1080/09620214.2021.2014933

Krosnick, J. A., & Presser, S. (2010). *Question and questionnaire design*. In P. V. Marsden & J. D. Wright (Eds.), Handbook of survey research (2nd ed., pp. 263–313). Emerald.

Lai, E. (2011). *Metacognition: A Literature Review*. Pearson Assessments Research Reports.

Lang, C. (2021). Extracurricular activities can play a central role in K-12 education. *Kappan, 102*(8), 14-19.

Leach, S. M., Immekus, J. C., French, B. F., & Hand, B. (2020). The factorial validity of the Cornell Critical Thinking Tests: A multi-analytic approach. *Thinking Skills and Creativity*, *37*, 100676. https://doi.org/10.1016/j.tsc.2020.100676

Ligita, T., Nurjannah, I., Wicking, K., Harvey, N., & Francis, K. (2022). From textual to visual: The use of concept mapping as an analytical tool in a grounded theory study. *Qualitative Research*, *22*(1), 126–142. https://doi.org/10.1177/1468794120965362

Lo, C. K., & Hew, K. F. (2020). A comparison of flipped learning with gamification, traditional learning, and online independent study: The effects on students' mathematics achievement and cognitive engagement. *Interactive Learning Environments*, *28*(4), 464–481. https://doi.org/10.1080/10494820.2018.1541910

Machado, C. T., & Carvalho, A. A. (2020). Concept Mapping: Benefits and Challenges in Higher Education. *The Journal of Continuing Higher Education*, *68*(1), 38–53. https://doi.org/10.1080/07377363.2020.1712579

Makhdum, F. N., Khanam, D. A., Faisal, A., & Sandhu, H. R. (2023). Impact Of Kahoot! On Students' Engagement And Learning Outcome At The Elementary Level In Pakistan: Their Perception Towards Kahoot! Assessment. *Journal of Positive School Psychology, 7*(1), 64-78. https://journalppw.com/index.php/jpsp/article/view/15017

Manzon, R. P. (2021). The Impact of Concept Mapping as a Teaching Strategy on the Students' Achievement and Interest in English. *EPRA International Journal of Research & Development (IJRD), 6*(8), 227–232. https://doi.org/10.36713/epra8348

McCosker, C., Renshaw, I., Russell, S., Polman, R., & Davids, K. (2021). The role of elite coaches' expertise in identifying key constraints on long jump performance: How practice task designs can enhance athlete self-regulation

in competition. *Qualitative Research in Sport, Exercise and Health*, *13*(2), 283–299. https://doi.org/10.1080/2159676X.2019.1687582

McMillan, J. H., & Schumacher, S. (2014). *Research in education: Evidence-based inquiry*. Pearson.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Merino-Armero, J. M., González-Calero, J. A., & Cózar-Gutiérrez, R. (2021). The effect of after-school extracurricular robotic classes on elementary students' computational thinking. *Interactive Learning Environments*, 1–12. https://doi.org/10.1080/10494820.2021.1946564

Merma-Molina, G., Gavilán-Martín, D., Baena-Morales, S., & Urrea-Solano, M. (2022). Critical Thinking and Effective Personality in the Framework of Education for Sustainable Development. *Education Sciences*, *12*(1), 28. https://doi.org/10.3390/educsci12010028

Moattari, M., Soleimani, S., Moghaddam, N. J., & Mehbodi, F. (2014). Clinical concept mapping: Does it improve discipline- based critical thinking of nursing students? *Iranian Journal of Nursing and Midwifery Research*, *19*(1), 70-76.

Mohammadi, F., Momennasab, M., Rostambeygi, P., Ghaderi, S., & Mousazadeh, S. (2019). The effect of education through conceptual mapping on critical thinking of nursing students. *Journal of Pakistan Medical Association, 69*(8), 1094-1098.

Mok, C. K. F., Whitehill, T. L., & Dodd, B. J. (2008). Problem-based learning, critical thinking and concept mapping in speech-language pathology education: A review. *International Journal of Speech-Language Pathology, 10*(6), 438–448. https://doi.org/10.1080/17549500802277492

Moroz, M., & Dunkley, D. M. (2019). Self-critical perfectionism, experiential avoidance, and depressive and anxious symptoms over two years: A three-wave longitudinal study. *Behaviour Research and Therapy*, *112*, 18–27. https://doi.org/10.1016/j.brat.2018.11.006

Mowling, C. M., & Sims, S. K. (2021). The Metacognition Journey: Strategies for Teacher Candidate Exploration of Self and Student Metacognition. *Strategies*, *34*(2), 13–23. https://doi.org/10.1080/08924562.2020.1867268

Moxon, J. (2022). Psychometric evaluation of abridged versions of the metacognitive awareness inventory in the Japanese population. *Current Psychology*. https://doi.org/10.1007/s12144-022-02763-2

Muncer, G., Higham, P. A., Gosling, C. J., Cortese, S., Wood-Downie, H., & Hadwin, J. A. (2021). A Meta-Analysis Investigating the Association Between Metacognition and Math Performance in Adolescence. *Educational Psychology Review*. https://doi.org/10.1007/s10648-021-09620-x

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equation in latent variable modeling with categorical and continuous outcomes*. Unpublished manuscript.

Muthén, L. K. (2013, January 28). *Multigroup ESEM* [3:21 pm]. Message posted to http://www.statmodel.com/discussion/messages/11/5581.html?1492019681.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2021). *Mplus (version 8.7)* [Computer software]. Muthén & Muthén.

Nawi, M.A.A, Ahmad, W.M.A.W, Rohim, R.A.A (2020). THE BEST WAY TO CHOOSING THE MULTIPLE COMPARISON TESTING FOR EQUAL VARIANCE AND UNEQUAL SAMPLE SIZE IN ONE WAY ANOVA. *International Journal of Public Health and Clinical Sciences, 7(*4), 1-13.

Ng, S.-Y., Cheung, K., & Cheng, H.-L. (2022). Critical Thinking Cognitive Skills and Their Associated Factors in Chinese Community College Students in Hong Kong. *Sustainability*, *14*(3), 1127. https://doi.org/10.3390/su14031127

Nikander, J., Tolvanen, A., Aunola, K., & Ryba, T. V. (2022). The role of individual and parental expectations in student-athletes' career adaptability profiles. *Psychology of Sport and Exercise*, *59*, 102127. https://doi.org/10.1016/j.psychsport.2021.102127

Ning, H. K. (2016). Examining heterogeneity in student metacognition: A factor mixture analysis. *Learning and Individual Differences*, *49*, 373–377. https://doi.org/10.1016/j.lindif.2016.06.004

Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in Psychology. *Review of General Psychology*, *23*(4), 403–424. https://doi.org/10.1177/1089268019883821

Novak, J. D. (2010). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations* (2nd ed). Routledge.

Nugroho, I. H., Susilaningsih, E., & Wijayati, N. (2019). Instrument Design to Measure the Critical Thinking Skill of Students that Participate in Chemistry National Science Olympiad. *Journal of Innovative Science Education, 8*(2), 147 - 152

Nwuba, I. S., Egwu, S. O., Fadekemi, O., & Osuafor, A. M. (2022). Secondary School Students' Critical Thinking Ability as Correlate of their Academic Achievement in Biology in Awka Education Zone, Nigeria. *Human Nature Journal of Social Sciences, 3*(4), 201-210. http://hnpublisher.com/ojs/index.php/HNJSS/article/view/74

Oderinu, O. H., Adegbulugbe, I. C., Orenuga, Omolola. O., & Butali, A. (2020). Comparison of students' perception of problem-based learning and traditional teaching method in a Nigerian dental school. *European Journal of Dental Education*, *24*(2), 207–212. https://doi.org/10.1111/eje.12486

Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning*, *13*(2), 179–212. https://doi.org/10.1007/s11409-018-9183-8

Omprakash, A., Kumar, A. P., Kuppusamy, M., Sathiyasekaran, B. W. C., Ravinder, T., & Ramaswamy, P. (2021). Validation of metacognitive awareness inventory from a private medical university in India. *Journal of Education and Health Promotion, 10*. 324. https://doi.org/10.4103/jehp.jehp_39_21

Oppong, E., Shore, B. M., & Muis, K. R. (2019). Clarifying the Connections Among Giftedness, Metacognition, Self-Regulation, and Self-Regulated Learning: Implications for Theory and Practice. *Gifted Child Quarterly*, *63*(2), 102–119. https://doi.org/10.1177/0016986218814008

Othman, M., Sahamid, H., Zulkefli, M. H., Hashim, R., & Mohamad, F. (2015). The Effects of Debate Competition on Critical Thinking among Malaysian

Second Language Learners. *Middle-East Journal of Scientific Research, 23*(4), 656-664. https://doi.org/10.5829/idosi.mejsr.2015.23.04.22001

Özçakır Sümen, Ö. (2021). The mediating role of metacognitive self-regulation skills in the relationship between problem-posing skills and mathematics achievement of primary pre-service teachers. *International Online Journal of Education and Teaching (IOJET), 8*(3). 2081-2096.

Paul, R. (2004). *The state of critical thinking today: as the organizer in developing blue prints for institutional change*. Retreived from: http://www.criticalthinking.org/professionalDev/the-state-cttoday.cfm.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, *61*(10), 991–996. https://doi.org/10.1016/j.jclinepi.2007.11.010

Pour, A. V., & Ghanizadeh, A. (2017). Validating the Persian version of metacognitive awareness inventory and scrutinizing the role of its components in IELTS academic Reading achievement. *Modern Journal of Language Teaching Methods, 7*(3), 46–63.

Powell, B. D., Oxley, M. S., Chen, K., Anksorus, H., Hubal, R., Persky, A. M., & Harris, S. (2021). A Concept Mapping Activity to Enhance Pharmacy Students' Metacognition and Comprehension of Fundamental Disease State Knowledge. *American Journal of Pharmaceutical Education*, *85*(5), 8266. https://doi.org/10.5688/ajpe8266

Prinz, A., Golke, S., & Wittwer, J. (2020). To What Extent Do Situation-Model-Approach Interventions Improve Relative Metacomprehension Accuracy? Meta-Analytic Insights. *Educational Psychology Review*, *32*(4), 917–949. https://doi.org/10.1007/s10648-020-09558-6

Purba, L. S. L. (2022). Analysis of critical thinking ability based on student gender through the implementation of independent curriculum in chemistry learning. *Jurnal Pendidikan Kimia*, *14*(3), 187–192. https://doi.org/10.24114/jpkim.v14i3.40305

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Rao, G. S. R., & Jaiswal, A. K. (2020). Evaluation of psychometric adequacy of metacognitive awareness inventory in Indian sample. *Indian Journal of Positive Psychology, 11*(2), 64-70.

Rebholz, F., Golle, J., Tibus, M., Ruth-Herbein, E., Moeller, K., & Trautwein, U. (2022). Getting fit for the Mathematical Olympiad: Positive effects on achievement and motivation? *Zeitschrift Für Erziehungswissenschaft*, *25*(5), 1175–1198. https://doi.org/10.1007/s11618-022-01106-y

Ren, L., & Zhang, X. (2020). Antecedents and consequences of organized extracurricular activities among Chinese preschoolers in Hong Kong. *Learning and Instruction*, *65*, 101267. https://doi.org/10.1016/j.learninstruc.2019.101267

Renatovna, A. G., & Renatovna, A. S. (2021). Pedagogical and psychological conditions of preparing students for social relations on the basis of the

development of critical thinking. *Psychology and Education Journal*, *58*(2), 4889–4902. https://doi.org/10.17762/pae.v58i2.2886

Rif'at, M., Rahmawati, P., Riyadi, S., & Heriyanto, H. (2022). Students' empirical thinking in solving mathematics problems. *Jurnal Kajian Pembelajaran Matematika*, *6*(1), 1. https://doi.org/10.17977/um076v6i12022p1-10

Romanko, L. (2016). *The Role Of Concept Mapping In The Development Of Critical Thinking Skills In Student And Novice Nurses: A Quantitative Meta-Analysis* (Master's thesis, The University of British Columbia). https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0228162

Roshangar, F., Azar, E. F., Sarbakhsh, P., & Azarmi, R. (2020). The Effect of Case-Based Learning with or without Conceptual Mapping Method on Critical Thinking and Academic Self-Efficacy of Nursing Students. *Journal of Biochemical Technology, 11*(1), 37-44.

Rusdin, R., & Rusli, R. (2020). The Optimizing Of Student's Mathematical Thinking Skills Through Matematika Nalaria Realistik In Mathematics Olympiad Training At Madrasah Ibtidaiyah. *Jurnal Pendidikan Dasar Islam, 12*(1), 97-110. https://doi.org/10.14421/al-bidayah.v12i1.305

Saleh, S. E. (2019). Critical Thinking as a 21st Century Skill: Conceptions, Implementation and Challenges in the EFL Classroom. *European Journal of Foreign Language Teaching, 4*(1), 1-16. https://doi.org/10.5281/ZENODO.2542838

Salmeen, A., Alkhaldi, N., Alshaber, R., & Majrashi, T. (2019). *Extracurricular Activities and Student Performance at Jubail University College*. *7*(1), 55-61.

Saud, W. I. (2020). The Relationship between Critical Thinking and Translation Ability of EFL Undergraduate Students. *International Journal of Social Sciences and Educational Studies*, *7*(3), 19-28. https://doi.org/10.23918/ijsses.v7i3p19

Schiefer, J., Golle, J., Tibus, M., Herbein, E., Gindele, V., Trautwein, U., & Oschatz, K. (2020). Effects of an extracurricular science intervention on elementary school children's epistemic beliefs: A randomized controlled trial. *British Journal of Educational Psychology, 90*(2), 382–402. https://doi.org/10.1111/bjep.12301

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning, 4*, 33–45. https://doi.org/10.1007/s11409-008-9031-3

Schraw, G., & Dennison, R. S. (1994). Asessing Metacognitive Awareness. *Contemporary Educational Psychology, 19*(4), 460-475. https://doi.org/10.1006/ceps.1994.1033

Schwarz, N., Jalbert, M., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review*, *4*(1), 4–25. https://doi.org/10.1002/arcp.1067

Shahzadi, U., Nimmi, & Khan, I. (2020). Exploring the Relationship between Critical Thinking Skills and Academic Achievement. *Sir Syed Journal of Education & Social Research, 3*(1), 236–242. https://doi.org/10.36902/sjesr-vol3-iss1-2020(236-242)

Siang, L. K., & Lan, N. P. (2011). Examining and comparing the factorial validity of the construct of metacognitive awareness across two grade levels. *1st*

International Conference on World Class Education. https://doi.org/10.13140/2.1.1211.4729

Silva, H., Lopes, J., Dominguez, C., & Morais, E. (2022). Lecture, Cooperative Learning and Concept Mapping: Any Differences on Critical and Creative Thinking Development? *International Journal of Instruction*, *15*(1), 765–780. https://doi.org/10.29333/iji.2022.15144a

Smith, A. K., Black, S., & Hooper, L. M. (2020). Metacognitive Knowledge, Skills, and Awareness: A Possible Solution to Enhancing Academic Achievement in African American Adolescents. *Urban Education, 55*(4), 625–639. https://doi.org/10.1177/0042085917714511

Song, Y., Lee, Y., & Lee, J. (2022). Mediating effects of self-directed learning on the relationship between critical thinking and problem-solving in student nurses attending online classes: A cross-sectional descriptive study. *Nurse Education Today, 109*. 105227. https://doi.org/10.1016/j.nedt.2021.105227

Souhila, D. H. (2022). Anticipating Students' Academic Achievements through Metacognitive Awareness and Learning Strategies Case Study of 3rd Year Students at the English Department Batna 2 University. *El-ihyaa journal, 22*(30), 1133 – 1348

Stevenson, M. P., Hartmeyer, R., & Bentsen, P. (2017). Systematically reviewing the potential of concept mapping technologies to promote self-regulated learning in primary and secondary science education. *Educational Research Review*, *21*, 1–16. https://doi.org/10.1016/j.edurev.2017.02.002

Strods, G., & Strode, A. (2018). A Critical Thinking Approach for Implementing Interdisciplinarity in Career Guidance for Secondary School Students. *CBU International Conference Proceedings*, *6*, 798–803. https://doi.org/10.12955/cbup.v6.1251

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education/Allyn & Bacon

Teo, T., & Lee, C. B. (2012). Assessing the Factorial Validity of the Metacognitive Awareness Inventory (MAI) in an Asian Country: A Confirmatory Factor Analysis. *The International Journal of Educational and Psychological Assessment, 10*(2), 92-103.

Tohir, M. (2019). Students' Creative Thinking Skills in Solving Mathematics Olympiad Problems Based on Metacognition Levels. *Alifmatika: Journal of Mathematics Education and Learning, 1*(1), 1-14. https://doi.org/10.31219/osf.io/uz89f

Tohir, M., & Muhasshanah, M. (2021). Mathematical Issues in Two-Dimensional Arithmetic for Analyze Students' Metacognition and Creative Thinking Skills. *Alifmatika: Jurnal Pendidikan dan Pembelajaran Matematika, 3*(2), 170-183. https://doi.org/10.35316/alifmatika.2021.v3i2.170-183

Toker, S., & Akbay, T. (2022). A comparison of recursive and nonrecursive models of attitude towards problem-based learning, disposition to critical thinking, and creative thinking in a computer literacy course for preservice teachers. *Education and Information Technologies*, *27*(5), 6715–6751. https://doi.org/10.1007/s10639-022-10906-y

Tseng, S.-S. (2019). Using Concept Mapping Activities to Enhance Students' Critical Thinking Skills at a High School in Taiwan. *The Asia-Pacific Education Researcher*, *29*(3), 249–256. https://doi.org/10.1007/s40299-019-00474-0

Utomo, D. P. (2018). An Analysis on Creative Thinking Skill on Algebra Materials of Students in Regular, Acceleration, and Olympiad Classes. *Advances in Social Science, Education and Humanities Research, 231*, 109-112.

Utomo, D. P., Putri, A. M. J., & Santoso T. (2022). Student's Critical Thinking Ability to Solve Problems HOTS in Regular, Acceleration, and Olympics Class Programs. *Jurnal Didaktik Matematika*, *9*(1), https://doi.org/10.24815/jdm.v9i1.23242

van Dongen, K., Schraagen, J. M., & Eikelboom, A. (2005). Supporting Decision Making by a Critical Thinking Tool. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 517-521.

Vandenberg, R. J. (2006). Introduction: statistical and methodological myths and urban legends. *Organizational Research Methods, 9*(2),194–201. https://doi.org/10.1177/1094428105285506

Verburgh, A., François, S., Elen, J., & Janssen, R. (2013). The Assessment of Critical Thinking Critically Assessed in Higher Education: A Validation Study of the CCTT and the HCTA. *Education Research International*, 1–13. https://doi.org/10.1155/2013/198920

Vettori, G., Vezzani, C., Bigozzi, L., & Pinto, G. (2018). The Mediating Role of Conceptions of Learning in the Relationship Between Metacognitive Skills/Strategies and Academic Outcomes Among Middle-School Students. *Frontiers in Psychology*, *9*, 1985. https://doi.org/10.3389/fpsyg.2018.01985

Vianin, P. (2016). Computerized Exercises to Promote Transfer of Cognitive Skills to Everyday life. *Perspective, 7*, 1-5. https://doi.org/10.3389/fpsyt.2016.00056

Wahidin, D., & Romli, L. A. M. (2020). Students Critical Thinking Development in National Sciences and Mathematics Competition in Indonesia: A Descriptive Study. *Jurnal Pendidikan IPA Indonesia*, *9*(1), 106–116. https://doi.org/10.15294/jpii.v9i1.22240

Wakelin, K. E., Perman, G., & Simonds, L. M. (2022). Effectiveness of self-compassion-related interventions for reducing self-criticism: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, *29*(1), 1–25. https://doi.org/10.1002/cpp.2586

Wang, M., Wu, B., Kirschner, P. A., & Michael Spector, J. (2018). Using cognitive mapping to foster deeper learning with complex problems in a computer-based environment. *Computers in Human Behavior*, *87*, 450–458. https://doi.org/10.1016/j.chb.2018.01.024

Wang, X., Wallace, M. P., & Wang, Q. (2017). Rewarded and unrewarded competition in a CSCL environment: A coopetition design with a social cognitive perspective using PLS-SEM analyses. *Computers in Human Behavior*, *72*, 140–151. https://doi.org/10.1016/j.chb.2017.02.045

Wiley, J., Griffin, T. D., Jaeger, A. J., Jarosz, A. F., Cushen, P. J., & Thiede, K. W. (2016). Improving metacomprehension accuracy in an undergraduate course context. *Journal of Experimental Psychology: Applied*, *22*(4), 393–405. https://doi.org/10.1037/xap0000096

Xue, X., Wang, Y., Li, H., Gao, J., & Si, J. (2021). The association between mathematical attitudes, academic procrastination and mathematical achievement among primary school students: The moderating effect of mathematical metacognition. *Current Psychology*. https://doi.org/10.1007/s12144-021-02133-4

Yanchar, S. C., Slife, B. D., & Warne, R. (2008). Critical Thinking as Disciplinary Practice. *Review of General Psychology, 12*(3), 265–281. https://doi.org/10.1037/1089-2680.12.3.265

Yang, C., Zhao, W., Yuan, B., Luo, L., & Shanks, D. R. (2022). Mind the Gap between Comprehension and Metacomprehension: Meta-Analysis of Metacomprehension Accuracy and Intervention Effectiveness. *Review of Educational Research, 0*(0). https://doi.org/10.3102/00346543221094083

Yang, Q.-F., Chang, S.-C., Hwang, G.-J., & Zou, D. (2020). Balancing cognitive complexity and gaming level: Effects of a cognitive complexity-based competition game on EFL students' English vocabulary learning performance, anxiety and behaviors. *Computers & Education*, *148*, 103808. https://doi.org/10.1016/j.compedu.2020.103808

Yennita, Y., & Zukmadini, A. Y. (2021). Problem-based learning (PBL) and blended learning in improving critical thinking skills and student learning activities in biochemistry courses. *Journal of Physics: Conference Series*, *1731*, 012007. https://doi.org/10.1088/1742-6596/1731/1/012007

Yuan, R., Liao, W., Wang, Z., Kong, J., & Zhang, Y. (2022). How do English-as-a-foreign-language (EFL) teachers perceive and engage with critical thinking: A systematic review from 2010 to 2020. *Thinking Skills and Creativity, 43*, 101002. https://doi.org/10.1016/j.tsc.2022.101002

Yue, M., Zhang, M., Zhang, C., & Jin, C. (2017). The effectiveness of concept mapping on development of critical thinking in nursing education: A systematic review and meta-analysis. *Nurse Education Today*, *52*, 87–94. https://doi.org/10.1016/j.nedt.2017.02.018

Zhang, Q. (Jackie), & Wang, L. (Peggy). (2014). Aggregating and Testing Intra-Individual Correlations: Methods and Comparisons. *Multivariate Behavioral Research*, *49*(2), 130–148. https://doi.org/10.1080/00273171.2013.870877

Zhao, N., Teng, X., Li, W., Li, Y., Wang, S., Wen, H., & Yi, M. (2019). A path model for metacognition and its relation to problem-solving strategies and achievement for different tasks. *ZDM*, *51*(4), 641–653. https://doi.org/10.1007/s11858-019-01067-3

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist, 25*, 3-17. https://doi.org/10.1207/s15326985ep2501_2

Zubaidah, S., Mahanal, S., Rosyida, F., & Kurniawati, Z. L. (2018). Using remap-TmPS learning to improve low-ability students' critical thinking skills. *Asia-Pacific Forum on Science Learning and Teaching, 19*(1), Article 11.

Zubova, S. P., Lysogorova, L. V., Kochetova, N. G., & Fedorova, T. V. (2021). Olympiad potential for identifying mathematical giftedness in elementary schoolers. *SHS Web of Conferences*, *117*, 02005. https://doi.org/10.1051/shsconf/202111702005

Zwaal, W., & Otting, H. (2012). The Impact of Concept Mapping on the Process of Problem-based Learning. *Interdisciplinary Journal of Problem-Based Learning*, *6*(1), Article 7, 104-128. https://doi.org/10.7771/1541-5015.1314