Macrothink Institute™

International Journal of Linguistics
ISSN 1948-5425
2024, Vol. 16, No. 7

Testing and Developing Critical Thinking in an ICLHE Context: Correlations Between Language Competence and Critical Thinking Skills

Christopher Liam McDonnell

Dipartimento di Scienze Umanistiche, Sociali e delle Imprese Culturali

Università di Parma

Strada Massimo D'Azeglio 85, 43125 Parma, Italy

Marco Mezzadri (Corresponding author)

Dipartimento di Scienze Umanistiche, Sociali e delle Imprese Culturali

Università di Parma

Strada Massimo D'Azeglio 85, 43125 Parma, Italy

E-mail: marco.mezzadri@unipr.it

Received: November 22, 2024 Accepted: December 23, 2024 Published: December 31, 2024

doi:10.5296/ijl.v16i7.22569 URL: https://doi.org/10.5296/ijl.v16i7.22569

Abstract

This contribution aims to investigate the potential correlations between performance on critical thinking (CT) assessments and the development of academic language competences in a vehicular language. The context is a preparatory programme (Foundation Year) at the University of Parma for international students who wish to continue their studies in Italy. Specifically, the participants in this study were divided into two subgroups: the first composed of students from Myanmar and the second consisting of students whose first or additional language is from the Slavic language family. Initially, students' gains on critical thinking assessments after explicit asynchronous critical thinking training were investigated and matched with their performance on a test for Italian as the vehicular language of the Foundation Year (FY) programme. Successively, the correlations between these two datasets were analyzed. The results of the analysis show that, although the CT instruction was effective and led to gains, a clear correlation between CT and performance on the language assessment was not found. Nonetheless, some tendencies as regards the overall development

of the students' language competences were noted, which testify to a possible link between CT and receptive language skills, despite the distance between students' L1 and the vehicular language. Further investigation should be devoted to analyzing the intersections between CT and a pervasive emphasis on thinking skills throughout the syllabus of the programme. Thus, a follow-up study has been planned which aims to consider more extensively the development of English language competences for academic purposes within the same educational context.

Keywords: Critical thinking, English medium instruction, Language proficiency

1. Introduction

This work stems from the training experience and the subsequent research path developed within the Foundation Year (FY) programme at the University of Parma. This is a one-year preparatory course aimed at future enrollment in a degree in either Italian or English. The students, from a variety of countries and diverse educational, cultural and linguistic backgrounds, find themselves in a kind of middle ground between school education and university, and are required to rapidly develop an appropriate level of language proficiency in a vehicular language and the study skills necessary for academic success. We believe that this is precisely the characteristic that sets this contribution apart from others, which are predominantly conducted at a stage where the respondents of the research have already been part of the academic context for some time (see Rashid & Hashim, 2008; Sanavi & Tarighat, 2014; Tabrizi & Jafari, 2015; Abduljabbar 2019) The languages used in the programme are Italian, with the goal of reaching at least a B2 proficiency level, as required by Italian law for university enrollment, or English and Italian for students who plan to enroll in an academic programme where English is used as the medium of instruction. In both cases, the focus of the curriculum concerning language training is on language for academic purposes. The entry-level language proficiency required for English is a minimum of B1. As for Italian, most students start with an initial or elementary level of competence.

The development of proficiency in a vehicular language to this level within a short period of time, alongside the reinforcement of the cognitive abilities necessary for academic success such as critical thinking (CT), is no easy feat. As a number of studies have demonstrated that CT skills correlate positively with academic success (Bala et al., 2022; Greatrix et al., 2021; O'Hare & McGuinness, 2009; Ren et al., 2020; Ross et al., 2013), this paper will attempt to investigate the potential correlations between students' development of CT and academic language proficiency in the vehicular language of the FY programme.

1.1 Studying in a Foreign Language at University: The Role of Soft Skills

To better define the context in which the research was conducted, we believe it is important to reflect on the characteristics of our specific context in relation to other areas where language is used for academic purposes, such as in English-Medium Instruction (EMI) or Content and Language Integrated Learning (CLIL) programmes. At the core lies the longstanding issue of the relationship between the acquisition of the language which is the medium of instruction on one hand, and the learning of disciplinary content on the other.

The context in which the research was conducted is partially defined by EMI dynamics, but differs from them in terms of the use of the vehicular language, which is partly English and partly Italian. It also differs in the greater level of integration between subject learning, through one or the other of the two languages depending on the discipline, and language learning for academic purposes, characterized by strong interactions with cognitive aspects. For this reason, we believe that investigating the development of soft skills such as CT in this context is of particular interest. From a linguistic standpoint, our context can be defined using the definition of EMI provided by Macaro et al. (2018, p. 37): "The use of the English language to teach academic subjects (other than English itself) in countries or jurisdictions where the first language of the majority of the population is not English".

Despite the coexistence of two vehicular languages in our project, we have adopted this definition because it seems to encompass everything needed to clear up any potential misunderstandings: there is no overlap between EMI and CLIL. In fact, in their definition, Macaro et al. (2018) clearly express the fundamental concept: EMI is not intended as a methodology for teaching the language, but rather defines the use of English for academic purposes to teach a subject. For further exploration of the distinction between EMI and CLIL, see Di Sabato and Macaro (2018, pp. 22-25).

CLIL, as a widely used teaching approach that deliberately focuses on the integrated learning of language and subject content, serves as the point of comparison with EMI. We believe this distinction is particularly relevant in educational contexts, at least in many European countries. For some authors, CLIL also represents the approach that differentiates pre-university study contexts from university-level ones. According to Lasagabaster (2022, p. 3):

> the lack of integration between language and content on the part of EMI university lecturers has led many researchers to disregard the CLIL acronym and use the label EMI at tertiary level. Aguilar (2017) points out that content is the priority in EMI and, although incidental language learning is expected (Banegas & Manzur Busleimán, 2021), there are usually no clear-cut and explicit language objectives.

Our context, on the other hand, presents a different situation, which can be defined, as mentioned above, using the metaphor of a "middle ground". This condition is ideal for us, as it allows us to develop and test our beliefs regarding the medium of instruction and the learning contexts in which a foreign language is used as a vehicle for instruction (see Mezzadri, 2017).

At the core is the belief that in a learning environment where a foreign language is used as a medium, the processes of language learning and subject-content learning coexist, and they are accompanied by the development of transversal skills related to cognitive, affective and relational spheres.

Moreover, in our "middle ground", learning occurs in a multilingual manner, and from this perspective, the conceptual framework of CLIL appears more coherent. In fact, according to

Dearden (2015, p. 4), "whereas CLIL is contextually situated (with its origins in the European ideal of plurilingual competence for EU citizens), EMI has no specific contextual origin".

Dearden further adds (2015, p. 4) that "Whereas CLIL does not mention which second, additional or foreign language (L2) academic subjects are to be studied in, EMI makes it quite clear that the language of education is English […]".

Despite the strong emphasis placed in the FY programme on language acquisition for academic purposes, the ultimate goal is to enable students to pursue educational careers in a vehicular language different from their mother tongue, thus allowing them to study academic subjects effectively. Consequently, the framework that seems to best represent our context is that of ICLHE (Integrating Content and Language in Higher Education), which extends a CLIL approach to the university level. As Smit states (2023, p.501), CLIL and ICLHE "focus on the pedagogical integration of content and target language […] in the teaching and learning of academic subjects other than the target language itself by teachers and students who are mainly L2 speakers of that target language." This implies that both teachers and students "need to develop a deeper understanding of the cognitive dimensions and methodological issues at play in ICLHE" (Lyster in Valcke & Wilkinson, 2017, p. 12). For an initial study on ICLHE in the Italian context, see Costa (2012).

Having clarified these aspects, the present contribution highlights the initial results of our research, which testify to the possibility of further investigation on the topic, especially as regards EMI within the FY programme. Our initial results will be presented in the Results and Discussion section.

1.2 Critical Thinking

Critical thinking (CT) is a crucial transversal skill in education (Liu et al., 2014; Pithers & Soden, 2010), which is essential for equipping students to face the challenges of modern societies (Facione, 2011). A survey conducted by the Association of American Colleges and Universities (AACU, 2011) revealed that 95% of researchers from 433 HE institutions in the USA indicated the development of CT as a key objective for all faculties. This tendency is reflected in the criteria used by the OECD in their Assessment of Higher Education Learning Outcomes, which place great value on CT (Tremblay et al., 2012). Employers also rank CT among the top five skills needed for success (Casner-Lotto & Barrington, 2006). However, a recent poll (Finley, 2021) showed that only 39% of U.S. executives believed graduates were well-prepared in CT, a concern supported by other studies (Arum et al., 2011; Willingham, 2008). Arum and Roska's longitudinal study assessed 2,322 U.S. students' progress using the Collegiate Learning Assessment, revealing "disturbingly low" gains in CT, complex reasoning, and writing (Arum et al., 2011).

However, the literature on CT has often struggled to provide a clear definition of this construct and how it can be developed. This debate is long-standing, with researchers such as Paul (2011) and Walters (1994) identifying three "waves of thought" which can be used to categorise the various definitions of CT proposed thus far.

The first or "philosophical" perspective on this construct dates to the second half of the 20th

century, and was characterised by a focus on defining the cognitive skills and dispositions involved in CT. The origins can be found in the work of Dewey (1910; 1933) and the Progressive Education Association's *Eight Year Study* (Aikin, 1942), whose participants had expressed concerns with the ability of school curricula to prepare students to be autonomous members of modern democratic societies. As a consequence, Glaser and later scholars such as Ennis saw the need to develop taxonomies of the skills and dispositions that students needed to possess to respond to these perceived difficulties, which could be targeted through instruction and assessed using standardised instruments. In Glaser's (1941) view, CT involved three aspects: a disposition towards thoughtful consideration of one's experiences, knowledge of the methods of logical reasoning, and some skill in using these methods. To assess the effectiveness of the teaching material based on these aspects, Glaser developed a standardised test which is still in use today: The Watson-Glaser Critical Thinking Appraisal.

The success of Glaser's work inspired other researchers, such as Ennis, to dedicate their attention to CT. In his initial conception, Ennis defined CT as the "correct assessing of statements" (Ennis, 1962, in Ennis, 1964, p. 599), which required students to be proficient in judging the reliability of sources and the logical solidity of arguments. This definition served as the framework for the *Cornell Critical Thinking Test* (Ennis & Millman, 1971). However, in contrast to Glaser and other previous thinkers, the notion of CT as a prerequisite for the development of a democratic society does not receive as much emphasis in Ennis. Furthermore, the importance of dispositional aspects plays a reduced role. Finally, there is a lack of focus on the consequences or actions that may arise from CT and no mention of CT being involved in problem solving or decision making. These aspects would later be addressed in his second conception (Ennis, 1987). Here, Ennis outlines that CT can be considered as "reasonable reflective thinking that is focused on deciding what to believe or do" (Ennis, 1987, p.10). Similarly to Glaser, Ennis believes that CT can be defined in terms of three macro-aspects (dispositions, abilities, and strategies and tactics), each of which could be targeted through instruction.

In 1987, the American Philosophical Association tasked Peter Facione to attempt to find a common ground among scholars, with the aim of producing a definitive account of the construct. The result of this request was the Delphi definition (Facione, 1990). In this document, considered by many to be the definitive account of CT (Davies, 2015), 46 experts agreed that CT involved two macro-components, cognitive skills and affective dispositions:

Table 1. Consensus List of Critical Thinking Skills and Sub-Skills (Facione, 1990, p. 12)

Skills	Sub-Skills
1. Interpretation	Categorization Decoding Significance Clarifying Meaning
2. Analysis	Examining Ideas Identifying Arguments

		Analyzing Arguments
3.	Evaluation	Assessing Claims Assessing Arguments
4.	Inference	Querying Evidence Conjecturing Alternatives Drawing Conclusions
5.	Explanation	Stating Results Justifying Procedures Presenting Arguments
6.	Self-Regulation	Self-examination Self-correction

Table 2. Affective Dispositions of Critical Thinking (Facione, 1990, p. 12)

Macro Area	Dispositions
2. Approaches to Life and Living in General	Inquisitiveness with regard to a wide range of issues. Concern to become and remain generally well-informed. Alertness to opportunities to use CT. Trust in the processes of reasoned inquiry. Self-confidence in one's own ability to reason. Open-mindedness regarding divergent world views. Flexibility in considering alternatives and opinions. Understanding of the opinions of other people. Fair-mindedness in appraising reasoning. Honesty in facing one's own biases, prejudices, stereotypes, egocentric or

	sociocentric tendencies. Prudence in suspending, making or altering judgments. Willingness to reconsider and review views where honest reflection suggests that change is warranted.
3. Approaches to Specific Issues, Questions or Problems	Clarity in stating the question or concern. Orderliness in working with complexity. Diligence in seeking relevant information. Reasonableness in selecting and applying criteria. Care in focusing attention on the concern at hand. Persistence though difficulties are encountered. Precision to the degree permitted by subject and circumstances.

In the 1980s and 1990s, the first wave's focus on democratic ideals and cognitive skills began to be criticised. Doubts began to arise as to whether its goals of encouraging a greater level of democratic participation among students and categorising the skills needed for success had actually been met, and as to their relevance in society with the advent of a new millennium. This issue was raised by scholars such as Thayer-Bacon (1998), who questioned the pertinence of previous conceptions developed in different historical contexts. Others such as McPeck (1985) highlighted the inherent complexity of the first wave's aims and criticised its tendency of considering CT as a trainable transversal skill, maintaining instead that it may manifest itself differently across disciplines. In addition, philosophers such as Barnett (1997) considered that traditional conceptions of CT were too narrow and did not pay enough attention to the need for critical action as a result of the thinking process.

More radical criticisms also began to surface from within the critical pedagogy movement, inspired by the work of Freire (Kaplan, 1991). For these scholars, the previous conceptions of CT were flawed, as their focus on abstract reasoning distracted from the true goal of any educational system: bringing about societal change and reducing oppression. Critics such as Giroux (1994) went further by arguing that traditional definitions showed a misunderstanding of the concept itself, in that their focus on cognitive skills ignored the influence of each

individual's belief system and values on the thinking process. Other scholars, such as Bailin (1995) and Thayer-Bacon (1998), added that this emphasis on logic at the expense of other forms of thinking was just one example of the many biases which were present in traditional definitions, which fostered an aggressive, individualistic approach among students that was at odds with their goal of developing a more equitable society.

Towards the turn of the millennium, these opposing positions led some observers to question the efficacy of both waves of thought. Burbules and Burk (1999) suggested that neither the "philosophical" or "critical pedagogy" perspectives had achieved their intended outcomes. Paul (2011) advocated for a third wave of thought, blending the rigour and focus on empirically provable results of the first wave with the inclusion of more varied forms of thought and the focus on critical action seen in the second.

Scholars such as Davies have attempted to provide new conceptions of CT which follow these indications. Concentrating specifically on the idea of CT in HE, Davies (2015) highlights how previous perspectives could induce institutions to focus on disparate aspects of this concept, from skills in logical reasoning, to developing *critical consciousness* in a Freirean sense, creating as a consequence a sense of confusion as to the true meaning of CT. Davies postulates that CT involves two separate dimensions: an individual dimension related to cognitive skills and dispositions, and a socio-cultural dimension related to the wider effects of CT on societal development. For Davies, the cognitive and propensity elements (dispositions) of CT fall into four categories:

Table 3. Cognitive elements of critical thinking (Davies, 2015, p.54)

Lower-level thinking skills	Higher-level thinking skills	Complex thinking skills	Thinking about thinking
Interpreting	Analyzing claims	Evaluating arguments	Metacognition
Identifying assumptions	Synthesizing claims	Reasoning verbally	Self-regulation
Asking questions for clarification	Predicting	Inference making	
		Problem solving	

Table 4. Propensity elements in critical thinking (Davies, 2015, p.58)

Dispositions arising in relation to self	Dispositions arising in relation to others	Dispositions arising in relation to the world	Other
Desire to be well-informed	Respect for alternative viewpoints	Interest	Mindfulness
Willingness to seek or be guided by reason	Open-mindedness	Inquisitiveness	Critical spiritedness
Tentativeness	Fair-mindedness	Seeing both sides of an issue	
Tolerance of ambiguity	Appreciation of individual differences		
Intellectual humility	Scepticism		
Intellectual courage			
Integrity			
Empathy			
Perseverance			
Holding ethical standards			

In this conception, the propensity elements motivate students to engage in CT. During this process, they utilise three levels of cognitive skills in order to complete their task. Finally, both the dispositions and cognitive skills are then refined through metacognition:

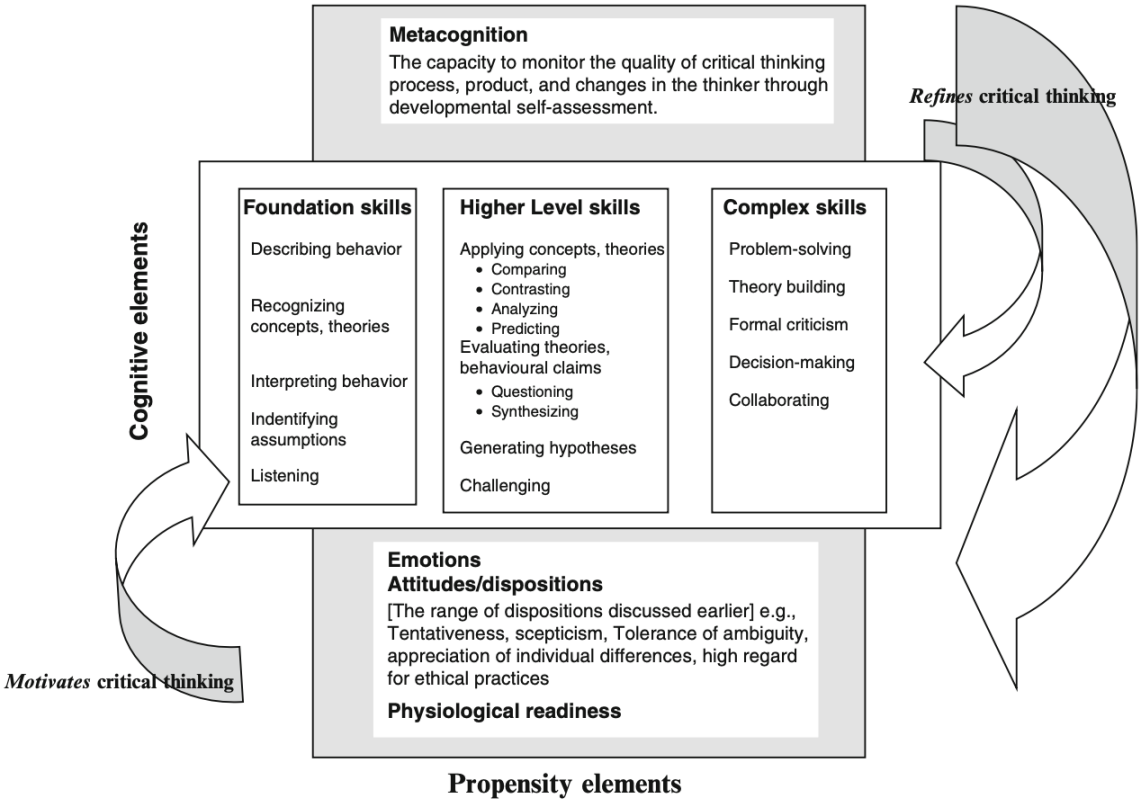

Figure 1. The cognitive and propensity influences on critical thinking (Halonen, 1995, in Davies, 2015, p. 62)

Davies suggests that institutions may use this model to implement CT instruction according to their individual aims, while respecting the common goal of increasing students' autonomy and freedom. In a later article, Davies (2019) proposes that this could be done in four ways: stand-alone CT instruction; "infusion" of explicit CT instruction into subject-matter classes using the specific content of these disciplines, "immersion" in which CT is believed to be implicitly developed through the standard syllabus, and a mixed-methods approach combining stand-alone courses with explicit instruction in subject-matter classes. It appears that the mixed-methods approach leads to greater gains in the development of CT skills and dispositions, while the immersion approach is less effective, as shown in a metanalysis of 117 studies conducted by Abrami et al. (2008). It is for this reason that the University of Parma has decided to adopt this approach within the FY programme.

1.3 CT in Language Teaching

Scientific literature has long investigated the centrality of CT in tertiary education (Phillips & Bond, 2004; Moore, 2004; Behar-Horenstein & Niu, 2011; Davies & Barnett, 2015), particularly within the Anglosphere.

Floyd (2011, p. 289) states that "Critical thinking (CT) […] is vital to success in tertiary level courses". Specifically addressing the role of CT in language education, it is clear that this is a particularly significant issue, although there is no clear consensus on the impact that studying in a second language combined with the application of CT may have. Floyd (2011, p. 289)

considers it "a significant issue, because not only is it important for educators to be aware of any CT limitations that students may have, but to understand more fully why limitations may exist". The topic, which has become central in language education (see Li, 2017; Liang & Fung, 2021), can rely on a school of thought that believes CT can facilitate the acquisition of language competence, particularly as regards receptive skills (Moore, 1995; Rashid & Hashim, 2008; Sanavi & Tarighat, 2014; Tabrizi & Jafari 2015; Abduljabbar 2019).

Lastly, we find it interesting to highlight an additional consideration: the available scientific contributions to date show a clear preference for this topic by scholars from Asian countries, possibly due to "the common perception of the Asian learner as lacking CT" (Floyd 2011, p. 289). In this regard, our study makes a contribution, thanks particularly to the inclusion of a significant subgroup of respondents in the sample: 42 students from Myanmar. Additionally, this study can count on a consolidated research path related to language for academic purposes at the University of Parma (see Mezzadri, 2017).

2. Methods

2.1 Context

The data presented in this study was gathered from students attending the FY programme at the University of Parma from September 2023-July 2024. This is a preparatory course addressed to foreign students interested in enrolling in a degree program at an Italian university or higher institution of music and art studies. The FY course is mainly devoted to the study of Italian, with a special emphasis on the language required for academic purposes, yet also maintains a strong focus on the skills required for future educational attainment within the areas of study chosen by students. For this reason, Italian language classes are taken alongside modules dedicated to CT, Intercultural Communication, and subject-specific courses offered by the various university departments. At the end of the programme, students take a test of Italian for academic purposes developed at the university, known as Italstudio. In line with the Common European Framework of Reference for Languages (CEFR) regarding the development of general competences (*savoir, savoir faire, savior être, savoir apprendre*), Italstudio emphasizes "learning to learn" as a fundamental element in the certification. To this end, it adopts a study skills syllabus organized by competence levels, and places continuous focus on balancing both inferential and non-inferential activities, as well as layering the tasks proposed. This second aspect is of particular importance, as Italstudio is committed to inclusive accessibility, minimizing any form of bias, whether cultural, linguistic, or related to the linguistic levels and awareness of test-takers in both their L1 and Italian L2. Therefore, no specific cultural or disciplinary knowledge is required to take the test. However, beside knowing Italian at a level corresponding to the parameters indicated for each level, candidates must possess the cognitive abilities and study skills developed through an adequate course of study in basic disciplines, either in their native language or in another foreign language.

Specifically, the data analysed below is taken from a pilot study of an elective CT course and the results of the final Italstudio test. Firstly, the effectiveness of CT instruction was assessed using a quasi-experimental research design with a pre- and post-test. Having obtained

positive results regarding the effect size of instruction as measured in Cohen's *d*, students' attainment on Italstudio was investigated, in the attempt to identify potential correlations between the level of acquisition of CT skills and academic language proficiency in a vehicular language.

The decision to allow the interaction of two coexisting linguistic environments in the FY programme is based on the conception of bilingual competence to which we refer. Specifically, it relies on the idea that linguistic skills are transferable between languages and that there is an interdependence between them, rather than a complete separation of skills. This approach finds its initial foundation in the work of Cummins (see Cummins, 1980 and 1984) and his reflections on the concepts of Common Underlying Proficiency (CUP) and Separate Underlying Proficiency (SUP). This is further supported by contributions from scientific research over the decades that has investigated the relationship between language skills and metalinguistic awareness (for a review, see Chłopek, 2018, pp. 24-27).

Essentially, the use of a vehicular language results in levels of performance closely linked to the level of linguistic-communicative competence in that language. However, in contexts where learners are required to use two vehicular languages in a comparable way - that is, to perform similar tasks in both - it becomes operationally indifferent which language is employed. This is, of course, contingent upon prior learning conditions and usage contexts: for example, the English of the students in the sample was mostly learned in a foreign language (FL) context, while Italian was acquired through immersion in an L2 context and was the dominant foreign language when the research was conducted. Obviously, there are other factors that play a major role, such as affective factors.

2.2 Materials

2.2.1 CT Course

The CT course analysed below comprised five modules, each of which focused on the cognitive skills and affective dispositions needed to face university admissions tests, such as the International Medical Admissions Test (IMAT) developed by Cambridge Assessment for English-language Medicine and Surgery courses in Italy. Each module contained a brief video lesson and a series of practical activities, on the following topics:

1. Introduction to CT and Logical Reasoning as an Academic Discipline
2. Assessing Claims
3. Assessing Evidence
4. Identifying and Analysing Argument
5. Interpreting and Evaluating Argument

Before accessing the course materials, participants took an initial assessment. Students then had one month in which to watch the video lessons and complete the provided activities. At the end of this period, they were then required to complete the final assessment. After analysis of the initial data from this pilot study, the decision was made to include an

expanded, in-person version of the course to other FY students. This course focused on the same topics.

2.2.2 CT Assessment

For the pre- and post-test, an ad-hoc instrument was created using authentic IMAT items from the seven different categories of question present on the test:

1. Summarising Conclusions
2. Drawing Conclusions
3. Identifying Assumptions
4. Assessing the Impact of Additional Evidence
5. Detecting Reasoning Errors
6. Identifying Parallel Reasoning
7. Applying Principles

Firstly, an analysis of the typology and quantity of CT items on IMAT past papers from 2011-2023 was conducted, in order to ensure similarity between the distribution of the items on the pre- and post-tests and previous editions of the IMAT. On the basis of this analysis, two instruments were created. Each instrument consisted of 30 questions, with 5 items respectively dedicated to summarising conclusions, drawing conclusions, identifying assumptions, assessing the impact of evidence, and detecting reasoning errors, 2 items dedicated to identifying parallel reasoning, and 3 items dedicated to applying principles. Once these test items were identified, they were then transformed into online quizzes using Microsoft Forms, which allowed for ease of administration and data analysis. A time limit of 60 minutes was set.

2.2.3 Italstudio

Italstudio is a certification of Italian for academic purposes which assesses the proficiency level of Italian as a second or foreign language (L2 or LS) in academic contexts. The test is divided into four sections: listening, reading, language use, written production. The listening section is composed of a pre-listening activity aimed at eliciting students' previous knowledge and orienting them as regards the content of the audio they will listen to. Subsequently, there is a global comprehension activity with both inferential and non-inferential items. The last task of this section consists in a note-taking activity. The reading section begins with a global reading activity, in which students must assign titles to the various paragraphs of the text. Then, students must complete a mind map using provided prompts based on the content of the reading comprehension text. Finally, students are presented with a typical cloze task which aims to measure their comprehension of a second academic text. The use of language section contains three tasks which aims to assess students' skills in three areas: morphosyntax, use of linking words, and punctuation. In conclusion, the written production section requires students firstly to describe and interpret a visual input of

an academic type (diagrams, tables, etc.), before writing an argumentative text based on the topic covered in the input (see Mezzadri, 2016).

2.3 Participants

2.3.1 CT Course

A total of 34 participants followed the online CT course, completing all of the required activities. 25 were female (73,5%) and 9 were male (26,5), with an average age of 20,5 years old. The vast majority of participants came from Myanmar (23 or 67.6%), followed by students from Russia (3), Turkey (3), Ukraine (2), Kuwait (1), France (1), and the USA (1). The data regarding the effectiveness of CT instruction is derived from this group. However, the data regarding performance on Italstudio includes both this pilot group and a group of students who attended the extended, in-person CT course.

2.3.2 Italstudio

As taking Italstudio is a requirement for completing the FY programme, the number of total participants is much larger than the one seen in the elective CT course. Data from 88 participants was analysed, who were divided into two subgroups: students from Myanmar (n=42) and students whose first or additional language is of Slavic origin (n=46).

3. Results and Discussion

In our considerations on these data, we will focus on two specific areas, CT and the Italstudio test, separately. Successively, we will illustrate the intersections between the two, which we believe could provide grounds for future research.

3.1 CT Results

Table 5. CT pre-test results

Participant	Summarising conclusions (/5)	Drawing Conclusions (/5)	Assumptions /5	Impact of evidence /5	Flaws /5	Parallel reasoning /2	Principles /3	TOTAL (/30)
1	2	2	2	1	2	0	1	**10**
2	3	2	2	0	0	0	0	**7**
3	3	3	2	0	1	1	1	**11**
4	4	2	3	1	1	0	2	**13**
5	2	2	2	1	1	0	1	**9**
6	3	3	2	3	4	1	1	**17**
7	4	3	2	2	0	2	2	**15**
8	3	0	2	1	2	2	1	**11**
9	2	1	2	4	0	1	2	**12**

10	1	3	2	2	1	1	2	**12**
11	3	4	3	3	3	2	2	**20**
12	3	1	2	1	1	0	0	**8**
13	2	2	3	1	2	0	0	**10**
14	3	2	2	3	1	0	0	**11**
15	2	4	3	4	4	2	3	**22**
16	2	0	0	0	1	0	2	**5**
17	3	1	2	1	1	1	1	**10**
18	2	2	2	2	2	0	1	**11**
19	0	2	2	1	1	0	1	**7**
20	3	3	4	1	1	1	2	**15**
21	4	2	1	3	0	0	0	**10**
22	3	1	2	4	2	0	0	**12**
23	2	2	3	1	0	0	0	**8**
24	4	2	1	2	2	1	1	**13**
25	2	3	2	1	2	0	0	**10**
26	3	1	2	0	1	1	1	**9**
27	1	3	3	1	2	2	1	**13**
28	2	4	1	3	3	1	1	**15**
29	2	2	1	4	1	0	0	**10**
30	2	2	3	0	0	1	1	**9**
31	3	1	1	1	0	0	1	**7**
32	3	3	3	4	2	1	2	**18**
33	2	1	1	2	1	1	1	**9**
34	4	5	3	4	2	2	2	**22**
Mean	2,56	2,18	2,09	1,82	1,38	0,71	1,06	11,79
SD	0,93	1,14	0,83	1,36	1,07	0,76	0,81	4,17

Table 6. CT post-test results

Participant	Summarising conclusions (/5)	Drawing Conclusions (/5)	Assumptions /5	Impact of evidence /5	Flaws /5	Parallel reasoning /2	Principles /3	TOTAL (/30)
1	4	3	2	3	1	1	0	**14**
2	2	2	3	2	4	0	3	**16**
3	1	4	0	0	1	0	3	**9**
4	4	5	2	1	3	2	3	**20**
5	4	1	2	2	2	1	2	**14**
6	4	5	3	2	3	2	2	**21**
7	3	2	3	5	2	1	2	**18**
8	3	3	2	1	3	2	2	**16**
9	3	2	0	3	4	1	1	**14**
10	4	4	4	0	1	1	2	**16**
11	4	3	2	3	3	2	3	**20**
12	3	4	4	2	4	0	3	**20**
13	2	1	1	1	2	1	0	**8**
14	3	2	2	2	2	1	1	**13**
15	3	5	3	4	4	1	2	**22**
16	4	2	3	2	0	1	0	**12**
17	2	2	2	2	4	0	0	**12**
18	5	3	2	0	3	2	2	**17**
19	2	2	0	1	1	1	1	**8**
20	2	5	2	2	5	1	2	**19**
21	2	3	2	1	4	1	1	**14**
22	5	4	3	1	3	1	1	**18**
23	1	2	0	2	0	0	0	**5**
24	3	4	2	2	1	1	3	**16**
25	3	5	3	3	2	0	0	**16**
26	5	3	3	0	2	0	0	**13**

27	1	1	0	1	2	2	0	**7**
28	3	3	4	3	3	2	3	**21**
29	3	2	1	1	3	1	1	**12**
30	2	3	1	1	4	0	2	**13**
31	4	4	3	0	1	0	1	**13**
32	4	5	5	2	4	2	2	**24**
33	3	3	2	2	0	1	1	**12**
34	2	3	4	5	4	2	3	**23**
Mean	3,03	3,09	2,21	1,82	2,50	1,00	1,53	15,18
SD	1,11	1,24	1,30	1,29	1,38	0,74	1,11	4,73

On average, participants scored 3.38 points higher on the post-test; however, it should be noted that this data contains positive outliers (Note 1) such as participants 2 (+9 points) and 12 (+12 points), and negative outliers (Note 2) such as participants 3 (-2), 13 (-2), 23 (-3), and 27 (-6). Excluding these outliers has the effect of reducing the variability in the dataset and leads to a slight increase in the average gain (+3.82) on post-test scores. The data outlined above allows for the calculation of the effect size of critical thinking instruction using Cohen's *d*. Regarding the complete data set, the effect size is $d = 0.759$, whereas excluding outliers this rises to $d = 0.916$. According to the interpretation guidelines expressed in Fritz et al. (2011) (Note 3), this signals a medium to large positive effect of instruction.

However, an aspect which must be mentioned is the relatively low overall score attained by students on both the pre-test and the post-test. We believe this could be attributed to a number of factors. Firstly, the items on the test were drawn from the Cambridge Assessment data bank and were originally developed for use on the BMAT in the UK. It follows that the target audience for these questions was intended to be students who aimed to study Medicine or Dentistry in an English-speaking country. While these items were used on the IMAT in Italy, it is possible that the level of linguistic complexity of the items was superior to the language competence of the students at the moment of assessment. Secondly, CT tasks such as the ones contained in the test items are not commonly covered in secondary school curricula. Therefore, these tasks were likely new for our participants. Thirdly, the duration of the explicit CT instruction was relatively brief (5 modules), although proven to be effective.

Upon a more detailed analysis of the data, we believe that there are a few aspects to consider which highlight possible correlations between thinking skills and the development of comprehension of the underlying logical structure of academic texts, such as the ones proposed by the Italstudio test. For instance, we can specifically appreciate the improvement demonstrated by members of the Myanmar subgroup in the final three tasks of the CT assessment: identifying flaws in reasoning, identifying parallel reasoning, and identifying the principles underlying an argument:

Table 7. Mean percentage increase on post-test by section

Subgroup	Summarising conclusions	Drawing Conclusions	Assumptions	Impact of evidence	Flaws	Parallel reasoning	Principles
Myanmar	16,66	36,36	16,66	-2,27	103,70	43,74	61,90
Control	22,22	50	-10,34	5,55	49,9	37,5	19,99

It is worthwhile noting that Myanmar students more than doubled (+103.7%) their score in tasks that require identification of flaws in reasoning. However, the most interesting gains, in our opinion, regard those seen in the parallel reasoning and principles sections (+43.74% and +61.9%), as these tasks require students to recognize the underlying logical structure of a text. These subskills are of paramount importance for Italstudio. In conclusion, we can argue that the participants of the study were able to transfer this set of acquired thinking skills to the reading comprehension task presented in Italstudio.

3.2 Italstudio

Table 8. Overall Mean Results

Subgroup	Listening /30	Reading /30	UoI /20	Writing /20	Total /100	SD
B2 Slavic-background (n=34)	22,57	24,90	13,51	13,71	74,69	6,46
B2 Myanmar (n=24)	17,67	26,65	15,08	14,25	73,65	7,29
B1 Slavic-background (n=11)	19,23	21,77	11,41	9,50	61,91	3,58
B1 Myanmar (n=17)	13,41	21,47	12,59	11,12	58,59	4,02

Table 9. Mean Results (no listening)

Subgroup	Reading /30	UoI /20	Writing /20	Total /70	SD
B2 Slavic-background (n=34)	24,90	13,51	13,71	52,12	5,45
B2 Myanmar (n=24)	26,65	15,08	14,25	55,98	5,83
B1 Slavic-background (n=11)	21,77	11,41	9,50	42,68	4,12
B1 Myanmar (n=17)	21,47	12,59	11,12	45,18	4,30

As illustrated above, Italstudio is characterized by a focus on language for study purposes and on the thinking and study skills required for academic contexts. Generally speaking, we can observe an overall high level of success among the participants of the study. With the exception of one student per subgroup, all participants attained at least B1, with many proceeding to obtain B2. However, the number of Myanmar students who obtained B2 is

notably inferior in percentage terms compared to the subgroup of students whose first or additional language is of Slavic origin (Myanmar = 57% B2/40.47% B1 vs. Slavic origin= 73% B2/ 23.9% B1). As regards the overall scores obtained by both groups, we can note a reflection of this tendency, with the average mark out of 100 being higher for students of a Slavic linguistic background in both the B1 and B2 groups (Myanmar =73.65% B2/58.59% B1 vs. Slavic origin 74.6% B2/61.69% B1). A potential justification could be the distance between the participants' L1 (Burmese/Slavic languages) and the vehicular language (Italian), seeing that the participants from both groups declared beginner or near-beginner levels of Italian language proficiency at the start of the FY programme.

Nonetheless, it is interesting to note that the average performance on the listening section was significantly lower among members of the Myanmar subgroup than that of members of the Slavic-background subgroup (Myanmar= 17.67 or 58.89% B2; 13.41 or 44.7% B1 vs Slavic-background= 22.57 or 75.25% B2; 19.23 or 61.4% B1). If we consider only the sections dedicated to written language competences (reading comprehension, use of Italian, and written production), the Myanmar subgroup outperformed the Slavic background subgroup (Myanmar=55.98/70 or 79.97% B2; 45.18/70 or 64.54% B1 vs. Slavic origin 51.12/70 or 74.58% B2; 42.68/70 or 60.97% B1).

4. Conclusions

In conclusion, we believe that the overall framework of the FY programme, based on intensive language instruction within a demanding context, such as those seen in ICLHE, that focuses strongly on the cognitive elements in language teaching/learning (e.g. the role of study and thinking skills) and the intersections between the different academic modules (e.g. CT and language courses), led to positive results for all students. This is true for both subgroups, regardless of the distance between participants' L1 and Italian. More specifically, however, we believe that the performance of the Myanmar subgroup in learning Italian for academic purposes is particularly significant, although the data currently in our possession does not allow us to highlight any clear correlations between the development of academic language skills and explicit CT instruction. After this pilot study, it is our intention to further investigate these aspects by subdividing our sample into two subgroups: one composed of participants who will receive explicit CT instruction, and a control group who will not receive formal training in CT. Furthermore, we intend to compare the results of this new study with the language competences demonstrated by students in English (the language of CT instruction within the FY programme). This will be possible due to the expansion of the FY programme, which is now more extensively offered in English. Finally, we believe that another area of research that could enrich our data is the investigation of the potential links between the metacognitive awareness of students and their performance on CT and academic language tasks, that could be measured through the use of tools such as the Metacognitive Awareness Inventory (Schraw & Dennison, 1994). An initial application of this tool was attempted during data collection for this study, with inconclusive yet potentially relevant results.

Acknowledgments

This study is promoted by the Inter-university Research Centre "L*in*E - Language *in* Education" (www.languageineducation.eu) within the scientific cluster "English-Medium Instruction in Higher Education: Needs Analysis and Training Initiatives in Italy". This publication is funded by the Strategic Plan 2022-2024 of the *Dipartimento di Lettere e Filosofia* of the University of Trento.

References

Abduljabbar, A. L. (2019). *The predictive utility of critical thinking skills on language proficiency in adult English as a second language learners*. Doctoral dissertation, Grand Canyon University. Retrieved from https://www.proquest.com/openview/554cbb406c4bb9872c60dc1778770585/1?pq-origsite=gscholar&cbl=18750&diss=y

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis. *Review of Educational Research, 78*(4), 1102-1134. https://doi.org/10.3102/0034654308326084

Aguilar, M. (2017). Engineering Lecturers' Views on CLIL and EMI. *International Journal of Bilingual Education and Bilingualism*. *20*(6), 722-35.

Aikin, W. M. (1942). *The Story of the Eight-Year Study With Conclusions And Recommendations Volume I*. Harper & Brothers. Retrieved from http://archive.org/details/storyoftheeighty009637mbp

Arum, R., Roksa, J., & Cho, E. (2011). *Improving Undergraduate Learning*. Social Science Research Council. Retrieved from https://files.eric.ed.gov/fulltext/ED514983.pdf

Association of American Colleges and Universities. (2011). The LEAP vision for learning: Outcomes, practices, impact, and employers' views. *Peer Rev.*, *13*(2), 34-34.

Bailin, S. (1995). Is Critical Thinking Biased? Clarifications and Implications. *Educational Theory*, *45*(2), 191-197. https://doi.org/10.1111/j.1741-5446.1995.00191.x

Bala, L., Pedder, S., Sam, A. H., & Brown, C. (2022). Assessing the predictive validity of the UCAT—A systematic review and narrative synthesis. *Medical Teacher, 44*(4), 401=409. https://doi.org/10.1080/0142159X.2021.1998401

Banegas, D. L., & Manzur Busleimán, G. (2021). EMI Materials in Online Initial English Language Teaching Education. In D. Lasagabaster, & A. Doiz (Eds.), *Language Use in English-Medium Instruction at University: International Perspectives on Teacher Practice* (pp. 100-25). New York: Routledge.

Barnett, R. (1997). *Higher Education: A Critical Business*. Buckingham: Open University Press.

Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning (TLC), 8*(2),

25-41.

Burbules, N. C., & Burk, R. (1999). Critical Thinking and Critical Pedagogy: Relations, Differences, and Limits. In T. Popkewitz, & L. Fendler (Eds.), *Critical Theories in Education* (0 ed., pp. 61-82). Routledge. https://doi.org/10.4324/9780203826256-11

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century US workforce.* Partnership for 21st Century Skills. Retrieved from https://files.eric.ed.gov/fulltext/ED519465.pdf

Chłopek, Z. (2018). "Is There a Relationship Between Language Competences and Metalinguistic Awareness?". In P. Romanowski (Ed.), *Current Research in Bilingualism and Bilingual Education* (pp. 22-39). Cham: Springer.

Costa, F. (2012). Focus on form in ICLHE lectures in Italy: Evidence from English-medium science lectures by native speakers of Italian. *alla Review*, *25*(1).

Cummins, J. (1980). The Cross-Lingual Dimensions of Language Proficiency: Implications for Bilingual Education and the Optimal Age Issue. *TESOL Quarterly*, *14*(2), 175-187.

Cummins, J. (1984). *Bilingualism and Special Education: Issues in Assessment and Pedagogy*. Clevedon: Multilingual Matters.

Davies, M. (2015). A Model of Critical Thinking in Higher Education. In M. B. Paulsen (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. 30, pp. 41-92). New York: Springer International Publishing. https://doi.org/10.1007/978-3-319-12835-1_2

Davies, M. (2019). New Directions in the Teaching of Critical Thinking. *Change*, *51*(5), 18-27. https://doi.org/10.1080/00091383.2019.1652063

Davies, M., & Barnett, R. (Eds.) (2015). *The Palgrave handbook of critical thinking in higher education*. New York: Palgrave Macmillan.

Dearden, J. (2015). *English as a Medium of Instruction. A growing global Phenomenon*. British Council, University of Oxford.

Dewey, J. (1910). *How We Think*. D.C. Boston: Heath & Co. Publishers.

Dewey, J. (1933). *How We Think*. D.C. Boston: Heath & Co. Publishers.

Di Sabato, B., & Macaro, E. (2018). Language and content intersections in CLIL and EMI. *RILA: Rassegna Italiana di Linguistica Applicata*, *2*(3), 21-32.

Ennis, R. H. (1964). A Definition of Critical Thinking. *The Reading Teacher*, *17*(8), 599-612.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron, & R. J. Sternberg (Eds.), *Teaching Thinking Skills: Theory and Practice* (pp. 9-26). New York: Freeman.

Ennis, R. H., & Millman, J. (1971). *Manual for Cornell Critical Thinking Test, Level X and Cornell Critical Thinking Test, Level Z*. University of Illinois: Critical Thinking Project.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of*

educational assessment and instruction (ED315423). American Philosophical Association.

Facione, P. A. (2011). Critical Thinking: What It Is and Why It Counts. *Insight Assessment*, *1*(1), 1-23.

Finley, A. (2021). *How College Contributes 'to' Workforce Success: Employer Views on What Matters Most.* Association of American Colleges and Universities. Retrieved from https://eric.ed.gov/?id=ED616977

Floyd, C. B. (2011). Critical thinking in a second language. *Higher Education Research & Development*, *30*(3), 289-302.

Fritz, C., Morris, P., & Richler, J. (2011). Effect Size Estimates: Current Use, Calculations, and Interpretation. *Journal of Experimental Psychology. General*, *141*, 2-18. https://doi.org/10.1037/a0024338

Giroux, H. A. (1994). Toward a pedagogy of critical thinking. In K. S. Walters (Ed.), *Re-Thinking Reason: New Perspectives on Critical Thinking* (pp. 199-204). New York: State University of New York Press.

Glaser, E. M. (1941). *An Experiment in Development of Critical Thinking*. Teachers College, New York: Columbia University.

Greatrix, R., Nicholson, S., & Anderson, S. (2021). Does the UKCAT predict performance in medical and dental school? A systematic review. *BMJ Open*, *11*(1), e040128. https://doi.org/10.1136/bmjopen-2020-040128

Halonen, J. S. (1995). Demystifying Critical Thinking. *Teaching of Psychology*, *22*(1), 75-81. https://doi.org/10.1207/s15328023top2201_23

Kaplan, L. D. (1991). Teaching Intellectual Autonomy: The Failure of the Critical Thinking Movement. *Educational Theory*, *41*(4), 361-370. https://doi.org/10.1111/j.1741-5446.1991.00361.x

Lasagabaster, D. (2022). *English-medium instruction in higher education*. Cambridge: Cambridge University Press.

Liang, W., & Fung, D. (2021). Fostering critical thinking in English-as-a-second-language classrooms: Challenges and opportunities. *Thinking Skills and Creativity*, *39*, 100769.

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment. *ETS Research Report Series*, *2014*(1), 1-23. https://doi.org/10.1002/ets2.12009

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of EMI in Higher Education. *Language Teaching*, *51*(1), 36-76.

McPeck, J. E. (1985). Critical Thinking and the 'Trivial Pursuit' Theory of Knowledge. *Teaching Philosophy*, *8*(4), 295-308. https://doi.org/10.5840/teachphil19858499

Mezzadri, M. (2016). *Studiare in italiano all'università. Prospettive e strumenti*. Torino: Bonacci Editore.

Mezzadri, M. (2017). *Testing Academic Language Proficiency*. Newcastle upon Tyne: Cambridge Scholars

Moore, R. (1995). *The relationship between critical thinking, global English proficiency, writing, and academic development for 60 Malaysian second language learners*. Doctoral dissertation, Indiana University. Retrieved from https://www.proquest.com/openview/5c193de3d94b48816631621b1020d38f/1?pq-origsite=gscholar&cbl=18750&diss=y

Moore, R. (1995). *The relationship between critical thinking, global English proficiency, writing, and academic development for 60 Malaysian second language learners*. Indiana University.

Moore, T. (2004). The critical thinking debate: How general are general thinking skills? *Higher Education Research & Development, 23*(1), 3-18.

Nikula, T., Dafouz, E., Moore, P., & Smit U. (Eds.) (2016). *Conceptualising Integration in CLIL and Multilingual Education*. Bristol: Multilingual Matters.

O'Hare, L., & McGuinness, C. (2009). Measuring Critical Thinking, Intelligence, and Academic Performance in Psychology Undergraduates. *The Irish Journal of Psychology, 30*(3-4), 123-131. https://doi.org/10.1080/03033910.2009.10446304

Paul, R. (2011). Critical Thinking Movement: 3 Waves. Retrieved from http://www.criticalthinking.org/pages/critical-thinking-movement-3-waves/856

Phillips, V., & Bond, C. (2004). Undergraduates' experience of critical thinking. *Higher Education Research & Development, 23*(3), 277-294.

Pithers, R. T., & Soden, R. (2010). Critical thinking in education: A review. *Educational Research, 42*(3), 237-249. https://doi.org/10.1080/001318800440579

Rashid, R. A., & Hashim, R. A. (2008). *The Relationship between Critical Thinking and Language Proficiency of Malaysian Undergraduates*, EDU-COM 2008 International Conference. Sustainability in Higher Education: Directions for Change, Edith Cowan University, Perth.

Ren, X., Tong, Y., Peng, P., & Wang, T. (2020). Critical thinking predicts academic performance beyond general cognitive ability: Evidence from adults and children. *Intelligence, 82*, 101487. https://doi.org/10.1016/j.intell.2020.101487

Ross, D., Loeffler, K., Schipper, S., Vandermeer, B., & Allan, G. M. (2013). Do Scores on Three Commonly Used Measures of Critical Thinking Correlate With Academic Success of Health Professions Trainees? A Systematic Review and Meta-analysis. *Academic Medicine, 88*(5), 724-734. https://doi.org/10.1097/ACM.0b013e31828b0823

Sanavi, R. V., & Tarighat, S. (2014). Critical Thinking and Speaking Proficiency: A Mixed-method Study. *Theory and Practice in Language Studies*. *4*(1), 79-87. https://doi.org/10.4304/tpls.4.1.1-7

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary*

educational psychology, *19*(4), 460-475.

Smit, U. (2023). English-medium instruction (EMI). *ELT Journal*, *77*(4), 499-503. https://doi.org/10.1093/elt/ccad018

Tabrizi, A. R. N., & Jafari, M. (2015). The Relationship among Critical Thinking, Self Efficacy, and Iranian EFL Learners' Reading Comprehension Ability with Different Proficiency Levels. *Academic Research International*, *6*(2), 412.

Thayer-Bacon, B. (1998). Transforming and Redescribing Critical Thinking: Constructive Thinking. *Studies in Philosophy and Education*, *17*(2), 123-148. https://doi.org/10.1023/A:1005166416808

Tremblay, K., Lalancette, D., & Roseveare, D. (2012). *Assessment of Higher Education Learning Outcomes. Feasibility Report: Design and Implementation*. OECD. Retrieved from https://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf

Valcke, J., & Wilkinson, R. (2017). *Integrating Content and Language in Higher Education: Perspectives on Professional Practice*. Frankfurt a.M: Peter Lang.

Walters, K. S. (1994). Introduction: Beyond Logicism in Critical Thinking. In K. S. Walters (Ed.), *Re-Thinking Reason: New Perspectives in Critical Thinking* (pp. 1-23). New York: State University of New York Press.

Notes

Note 1. Defined as a post-test difference of over +8.66, using the standard formula (mean + (1.5 x SD)).

Note 2. Defined as a post-test difference of under -1.9, using the standard formula (mean - (1.5 x SD)).

Note 3. Small effect size (*d = 0.2*); medium effect size (*d = 0.5*); large effect size (*d = 0.8*).

Copyrights