Received: 31 July 2023 | Accepted: 7 November 2023

DOI: 10.1002/rev3.3442

ORIGINAL ARTICLE

Review of Education BERA

A multi-country comparison of lower secondary students' critical thinking under the International Baccalaureate and national curricula

Samantha-Kaye Johnston[1] | **Joshua A. McGrane**[2,3] | **Mireia Vendrell-Morancho**[4] | **Therese N. Hopfenbeck**[2,3]

[1]Department of Computer Science, University of Oxford, Oxford, UK

[2]Assessment and Evaluation Research Centre (AERC), Faculty of Education, University of Melbourne, Melbourne, Victoria, Australia

[3]Kellogg College, University of Oxford, Oxford, UK

[4]Faculty of Education, Universidad Complutense de Madrid, Madrid, Spain

Correspondence
Samantha-Kaye Johnston, Department of Computer Science, University of Oxford, Oxford, UK.
Email: samantha-kaye.johnston@cs.ox.ac.uk

Funding information
International Baccalaureate Organisation

Abstract

Critical thinking is an essential skill for life-long learning, and, given its increasing importance as a graduate attribute, it is vital to evaluate how educational systems can best improve students' critical thinking through their curricula and classroom practices. This study evaluates the differences in the critical thinking skills of students in the International Baccalaureate (IB) Middle Years Programme (MYP), which explicitly emphasises critical thinking development in its curriculum and has been shown to facilitate classroom strategies to develop critical thinking, versus students enrolled in the national curricula in Australia, England and Norway. The study comprised 870 students across the MYP ($n = 386$) and non-MYP ($n = 484$) curricula in Grades 9 and 10 across 21 schools. Data were remotely collected on their critical thinking skills, and several other relevant cognitive, non-cognitive and background characteristics. A propensity score matching approach was used to match the MYP and non-MYP groups on these characteristics and their critical thinking skills were compared using weighted regression. Findings showed that, overall, MYP students possessed higher levels of critical thinking skills than their non-MYP peers with a moderate effect size ($\beta = 0.38$). This advantage also held at both grade levels and across Australian and English

students, with no difference for Norwegian students. Thus, the MYP appears to be a promising exemplar for enhancing critical thinking among secondary-aged students, although context needs to be considered. Based on these findings, recommendations of specific instructional strategies are offered, as well as future research to inform criticalthinking pedagogy.

KEYWORDS

assessment, critical thinking, International Baccalaureate Middle Years Programme (MYP), secondary school

Context and implications

Rationale for this study

Critical thinking, and its integration in curricula, has been researched for over 100 years, yet little research has robustly compared different curricula in fostering critical thinking.

Why the findings matter

Our findings matter as they point to an effective way of fostering students' critical thinking through an established curriculum at the secondary school level.

Implications for practitioners, policy makers, researchers and funders

Practitioners may consider enhancing their teaching of critical thinking by adopting inquiry-based methods, enabling students to apply their critical thinking across contexts and subjects. Additionally, policy makers should ensure that critical thinking is explicitly integrated into curricular goals and provide professional development opportunities for teachers in this area. Although explicit teaching is crucial, there is still a requirement for robust methodological designs that compare different approaches to teaching critical thinking, including explicit, immersion, infusion and mixed methods. Therefore, funders and researchers should evaluate the effectiveness of these methods across different educational curricula, grade levels and dimensions of critical thinking.

INTRODUCTION

Critical thinking has become a goal for several countries, and its teaching in the classroom is viewed as one of the most foundational competencies to develop among students (Meirbekov et al., 2022; Yuan et al., 2022). Common taxonomies of twenty-firstcentury skills typically list critical thinking as one of the top skills needed for students to thrive in a rapidly changing world (Dilekçi & Karatay, 2023; van Laar et al., 2020; Xu et al., 2023). While there are various definitions of critical thinking, Dewey (1909) viewed it as a reflective process in which one actively considers

evidence-based knowledge and draws warranted conclusions. Since Dewey's conceptualisation, other authors have defined critical thinking as the ability to consider assumptions and to employ objective, logical and well-reasoned arguments based on diverse perspectives. For example, Ennis & Weir (1985) built on Dewey's definition and defined critical thinking as 'reflective and rational thinking that focuses on deciding what to believe or do' (p. 45). Yet, with the increase in students engaging in digital environments, critical thinking has expanded to indicate an ability to identify misleading information and digital designs, including deep fakes (Diepeveen & Pinet, 2022). In the current study, we adopt a widely accepted definition derived from the Delphi Report commissioned by the American Philosophical Association in 1988 (Ennis, 1989, 1990): critical thinking is defined as 'purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based' (Facione, 1990, p. 3). Despite being unwieldy, this definition has become widely accepted and provided the framework for the most comprehensive meta-analysis regarding the instruction of critical thinking (Abrami et al., 2008). Moreover, this definition embodies the similarities identified across different conceptualisations of critical thinking regarding its reference to questioning assumptions, employing higher-level thinking and problem solving (Wallace & Jefferson, 2015).

The cultivation of critical thinking skills has several positive effects on a student's life, including increased academic success (Akpur, 2020; D'Alessio et al., 2019), greater resistance to misinformation (Joshi et al., 2022), better decision-making (Butler et al., 2017), access to future higher-quality employment opportunities (Goodsett, 2020; Indrašienė et al., 2021), and the ability to become informed, engaged and ethical citizens who contribute to the building and sustaining of a healthy and democratic society (Aktoprak & Hursen, 2022; Tiruneh et al., 2016). In today's information age, where vast amounts of information are accessible through online environments, the ability to think critically is more important than ever before. It is crucial to equip young people with the necessary skills to evaluate and interpret information effectively. Importantly, although parents and relatives play an essential role in encouraging children to appropriately use digital devices, they do not always succeed in maximising online opportunities while reducing potential risks (Livingstone et al., 2017; Rodríguez-de-Dios et al., 2018). Therefore, providing equitable opportunities to learn digital skills at school, including critical thinking, is not only beneficial to all students but could also help to mitigate some of the learning gaps because of the second digital divide among parents, including those from lower socioeconomic backgrounds (Hargittai et al., 2019). In the current study, we examine the development of critical thinking skills among lower secondary students[1] across different curricula in Australia, England and Norway: these countries were specifically selected due to their diverse educational systems, cultural contexts and their emphasis on critical thinking in their national curricula. Importantly, although each country has its own unique reality, we were confident in our ability to understand the local context. This confidence stems from our previous experiences living and working within these regions, or from maintaining professional connections with individuals from these countries. To clarify, we recognise the uniqueness of each country and the nuances that distinguish them from one another. However, our familiarity with these regions allowed us to effectively navigate their distinct contexts.

In Australia, critical thinking is considered a general capability, which is part of the broader critical and creative thinking capability, and it is taught using various strategies and approaches, such as inquiry-based learning, problem-based learning and using digital technologies and multimedia resources (Australian Curriculum Assessment and Reporting Authority [ACARA], n.d.). England's national curriculum does not specifically mention critical thinking. However, it emphasises critical thinking behaviours, such as analysing and evaluating information, arguments and evidence across different subjects (Department for Education [DfE], 2013a). By contrast, Norway places great emphasis on critical thinking, explicitly including it in the Core Curriculum's values and principles for primary and secondary education (Norwegian Ministry of Education

and Research, 2017). In addition, the Norwegian core curriculum (hereafter Norwegian national curriculum) also emphasises ethical awareness and encourages students to consider the ethical implications of their actions and decisions. In the next section, we delve into the national curriculum implemented in each of these three countries, with a focus on the instructional strategies employed to cultivate critical thinking. Effective pedagogical approaches that foster critical thinking in the classroom will be emphasised, before providing an overview of this study.

Differences in teaching critical thinking across different educational contexts

Critical thinking education is viewed as a foundational component of recent curricula changes (Ab Kadir, 2018; Lombardi et al., 2021). More importantly, international comparisons on approaches to teaching critical thinking in different contexts provide valuable insights for understanding best practices that improve higher order thinking skills, such as critical thinking (Abrami et al., 2015). The teaching of critical thinking is influenced by both the national curriculum and cultural context across various educational contexts. While we acknowledge that critical thinking instruction is an important pedagogical goal across many educational systems, we focus our discussion on Australia, England and Norway, given that our study was conducted in these three countries.

While Australia, England and Norway all recognise the significance of developing critical thinking, there exists at least three differences in their approaches to teaching and integrating critical thinking into their national curricula and supporting its growth (see Table 1; ACARA, n.d.; DfE, 2013a; Norwegian Ministry of Education and Research, 2017). The first relates to the level of specificity provided by each curriculum. The Australian curriculum defines critical thinking as a general capability and provides a comprehensive set of learning outcomes and a detailed learning continuum that outline the skills and knowledge that students should acquire as they progress through their education (ACARA, 2014). Given the complexity and ambiguity of critical thinking, it is beneficial to provide students with a clear and comprehensive definition of the concept and its relevance to their subjects to cultivate this skill effectively (Bezanilla et al., 2019; Manalo et al., 2015). Therefore, explicitly defining critical thinking and identifying its essential elements in the curriculum would benefit both educators and students in teaching and learning this vital skill. Moreover, since students tend to learn based on how they are evaluated, evaluation and assessment are integral to the development of critical thinking in educational contexts (Dwyer, 2017; Haynes et al., 2015). Thus, clearly stated goals and learning objectives for critical thinking, followed by the standards to achieve them, would be beneficial. In contrast, the English and Norwegian curricula offer more general guidance on critical thinking. While England's national curriculum makes few explicit references to critical thinking (the only subject that explicitly states critical thinking as a key goal is Citizenship), it promotes critical thinking behaviours, such as analysing and evaluating information, arguments and evidence across various subjects (DfE, 2013a). The Norwegian core curriculum includes critical thinking as a core value and principle for primary and secondary education (Norwegian Ministry of Education and Research, 2017), but it is less specific than the Australian curriculum in its approach. However, in the most recent update to the Norwegian curriculum (LK20), there is a focus on the core competencies that students should acquire at different stages alongside offering teachers the autonomy to define their own pedagogical approaches (Norwegian Directorate for Education and Training, 2020).

The second difference is the degree of integration of critical thinking across subject areas. In Australia, critical thinking is perceived as separate from the subject curriculum and taught as a general capability; it is recommended across all subject areas but, given that it is not precisely defined, it is often integrated in a piecemeal and uncoordinated manner

TABLE 1 Critical thinking (CT) across national curricula in Australia, England and Norway.

	Australia	England	Norway
Level of specificity of CT	• CT as a general capability with detailed learning continuum	• General guidance on CT	• CT is positioned as a core value and principle
Degree of integration of CT across different curriculum areas	• Piecemeal integration across all subject areas as a general capability	• Integrated into subject-specific learning objectives	• Integrated across all subject areas
Approaches to teaching CT	• Inquiry and problem-based approaches • Use of digital technologies • Emphasis on creativity in critical thinking	• Inquiry and problem-based approaches • Use of digital technologies • Emphasis on evaluation and analysis	• Inquiry and problem-based approaches • Use of digital technologies [not given the same level of emphasis as Australia & England] • Emphasis on evaluation and analysis

Note: The information presented above is primarily based on the following sources: (1) Australian Curriculum Authority [ACARA] (2014). *Assessment and Reporting Authority: Critical and creative thinking*. ACARA. (2) Department for Education [DfE] (2013b). *National curriculum in England: Framework document*. DfE. (3) Norwegian Ministry of Education and Research [NMER] (2017). *Core curriculum for primary, lower secondary and upper secondary education: Introduction*. NMER.

(Gilbert, 2019), while in Norway it is integrated into the national curriculum across all subject areas. In contrast, in England, it is integrated into subject-specific learning objectives. This difference may have implications for students' ability to develop critical thinking skills. For example, the focus on teaching critical thinking as a general capability in Australia may result in less thorough or systematic teaching compared to Norway, where it is integrated into the national curriculum (Dwyer, 2017; Paul & Elder, 2019). Consequently, Australian students may not develop the same level of critical thinking skills as students in Norway. On the other hand, the compartmentalised approach to teaching critical thinking in England, where it is integrated into subject-specific learning objectives, may hinder the development of the skill for students who struggle in those subjects. Moreover, the differences in teaching critical thinking across these educational contexts may have implications for the way students approach ethical issues. In Norway, where critical thinking is integrated across all subject areas, students may be more likely to approach ethical issues from a critical and independent perspective. In contrast, in Australia and England, where critical thinking is taught in a more compartmentalised way, students may be less likely to apply critical thinking skills to ethical issues outside of specific subject areas. In recent years, there has been a growing emphasis on the importance of cultivating critical thinking through real-life problems, social issues, and moral and controversial dilemmas (Santos et al., 2021; Velez & Power, 2020; Volman & ten Dam, 2015). It is important to note that critical thinking is not an innate ability, but rather a learned skill that requires explicit instruction and ongoing practice (Dwyer, 2017; Evangelisto, 2021; Zhang et al., 2022). As a result, educators must prioritise the development and application of critical thinking across different domains, including academia, personal life, civic engagement and work-related settings. To accomplish this objective, critical thinking should be integrated throughout the curriculum at the institutional level (Ennis, 2018). Additionally, at the subject level, providing students with opportunities to think broadly is essential. To achieve these goals, Halpern (2014) suggests arranging learning that is less reliant on content and more focused on cultivating skills and dispositions using a variety of instructional materials based on real-world contexts.

Finally, a noticeable difference among these educational contexts is their approach to teaching strategies. While all three prioritise inquiry-based and problem-based learning (Archila et al., 2022; Pnevmatikos et al., 2019), as well as the use of digital technologies to support critical thinking skills, the Australian curriculum stands out for its emphasis on creativity in critical thinking (ACARA, 2014). On the other hand, the English (DfE, 2013a) and Norwegian curricula (Norwegian Directorate for Education and Training, 2020) place more emphasis on the evaluation and analysis of information and arguments. Additionally, there are differences in the extent to which digital technologies are emphasised in teaching and learning. The Australian curriculum strongly emphasises the use of digital technologies to support critical thinking skills. In England and Norway, while digital technologies are recognised as important tools, they are not given the same level of emphasis. Familiarity with navigating digital technologies is important given the rapid changes in societal requirements; it is essential for students to cultivate their thinking and reasoning skills to effectively handle instances of misinformation and meet the demands of the future workforce (Goodsett, 2020; Musi et al., 2023).

Despite the above-mentioned differences in the approaches to teaching and integrating critical thinking in Australia, England and Norway, what is common across each context is that, alongside the national curricula, they also offer the International Baccalaureate (IB) Middle Years Programme (MYP). The IB has a fundamental mission of nurturing globally conscious individuals who acknowledge their shared humanity and the collective responsibility for the planet (IBO, 2018). To fulfil this mission, the IB framework aims at fostering the IB learner profile, which underscores the holistic development of attributes such as curiosity, open-mindedness, risk-taking, and balanced thinking (IBO, 2018). In addition to promoting cognitive growth, the IB programme places importance on students' social, emotional, and

physical well-being, instilling a deep sense of respect for oneself, others, and the world (IBO, 2018). Furthermore, the four distinct IB programmes (primary, middle school, diploma and career levels) offer comprehensive, interconnected, and conceptually driven curricula that actively encourage interdisciplinary learning within real-world contexts (IBO, 2018). Within the IB, the MYP stands out due to its distinctive emphasis on nurturing critical thinking skills, which is widely acknowledged as a fundamental component of the programme (Calnin et al., 2018; Dever, 2019). For example, it has a specific unit, Individuals and Societies, that is solely dedicated to fostering and assessing critical thinking. This commitment aligns with its global presence, evident in its implementation in 1859 schools across 127 countries as of November 2023, reflecting its widespread adoption, partly attributed to its potential in fostering critical thinking (Ateşkan et al., 2016; Holland, 2016; IBO, 2023; Robertson, 2011; Wright et al., 2016). Importantly, the MYP, tailored for students aged 11 to 16, occupies a unique position in the IB framework, bridging the gap between the Primary Years Programme (PYP) and the Diploma Programme (DP). Another distinctive feature of the MYP is its focus on six global contexts, which build upon and extend the PYP transdisciplinary themes (IBO, 2022). These global contexts encompass identities and relationships, personal and cultural expression, orientation in space and time, scientific and technical innovation, fairness and development, and globalisation and sustainability (IBO, 2018). While all IB programmes share a common commitment to international-mindedness and the development of the IB learner profile attributes, the MYP sets itself apart by emphasising these global contexts to provide a unique and developmentally appropriate educational experience (IBO, 2022).

To realise the IB's (and by extension the MYP's) vision of international mindedness, which supports critical thinking disposition (i.e., the willingness to think critically), the IB anchors its approach in repeated cycles of sustained inquiry, critical reflection, and principled action (Giancarlo et al., 2004; IBO, 2018). Principled action encapsulates the active measures integrated into the MYP to ensure students actively engage in critical thinking. It encompasses fostering an ethos of thinking and implementing active strategies within the programme (IBO, 2014a). Central to this ethos of thinking is the pivotal role of critical thinking, encompassing the analysis and evaluation of issues and ideas, including reasoning and problem-solving sub-skills (IBO, 2014a). The MYP also promotes an array of critical subskills, including communicative, creative, conceptual, metacognitive, media and information literacy, research, reflective, self-management, and social skills, enriching the development of critical thinking. Moreover, the MYP underscores the significance of a robust knowledge foundation for fostering clear and precise thinking about complex issues (IBO, 2014a). In the pursuit of active critical thinking skills within the MYP, various strategies and elements are integral. These include the integration of key concepts, related concepts, and global contexts to formulate meaningful statements of inquiry, thereby empowering students to develop their questions, deepen their understanding, and engage in meaningful reflection and responsible action. Multilingualism, interdisciplinary learning, a concept-driven curriculum, and specific academic subjects all contribute substantially to the cultivation of critical thinking skills. Additionally, the MYP places paramount importance on action, serving as both a strategy and outcome, thus encouraging students to apply their critical thinking skills in real-world experiences (IBO, 2014a).

A profound understanding of language is foundational for the development of critical thinking skills. It empowers students to decode and comprehend information effectively, evaluate evidence, construct counterarguments, and propose solutions (IBO, 2014b). The MYP curriculum prioritises language education to equip students with the requisite skills for critical engagement in contemporary societies. This goal is accomplished through a myriad of strategies, including language learning, mother-tongue development, language proficiency across subjects, and explicit instruction on language features (IBO, 2012, 2014b). Moreover, assessment within the MYP framework plays an indispensable role in nurturing

and incentivising the development of critical thinking abilities. Critical thinking is assessed comprehensively across the curriculum, with specific emphasis within distinct subject areas (e.g., within the Individual and Societies unit). The MYP's transparent assessment criteria guide educators in evaluating student performance across subject groups, with a sharp focus on knowledge dimensions and higher-order thinking skills. Stringent accountability measures, including internal standardisation and clear task expectations, aim to ensure the fidelity of assessment implementation. Central to this is that the MYP offers assessments such as the Personal Project, ePortfolios and onscreen examinations, all of which underscore higher-order thinking and problem-solving (IBO, 2014a, 2018).

Empirical evidence lends support to the assertion that the above-mentioned features of the MYP enhance students' critical thinking. For example, studies across various nations, including Australia, Turkey, Spain, the United Kingdom, the United Arab Emirates, the United States and several Asian countries, consistently reported a positive impact on students' critical thinking development within the MYP framework, as perceived by students, parents, teachers and school leaders (Aoki, 2018; Ateşkan et al., 2016; Australian Council for Educational Research [ACER], 2015; Beckwitt et al., 2015; Sizmur & Cunningham, 2013; Stevenson et al., 2017; Storz & Hoffman, 2018; Wade et al., 2016; Wolanin & Wade, 2012). Nevertheless, it is imperative to acknowledge that perceived benefits may not always precisely align with the actual impact of the MYP on students' critical thinking development. Consequently, further empirical research is indispensable to determine definitively whether the MYP genuinely outperforms other curricula in fostering critical thinking (Dickson et al., 2018; Wootten, 2019).

Current study

An examination of IB programmes revealed that few studies have employed methodological designs (e.g., accounting for confounding variables) that allow conclusions to be suggested about the causal role of IB participation. Moreover, it is surprising that although the MYP accounts for around 20% of authorised programmes globally, it is the subject of very little research in comparison with other IB programmes (Dickson et al., 2018; IBO, 2022). For example, Dickson et al. (2018) found that only 6% of articles have examined the role of the MYP in student learning outcomes compared with 76% for the Diploma Programme (DP) and 11% for its Primary Years Programme (PYP). This could be related to the nature of the DP, which prepares students for university entrance, and is therefore a significant concern for educators and policy makers. This focus on higher education often leads to more research aimed at understanding the programme's impact on college preparedness and success. By contrast, the MYP, while important, may not be as directly connected to college preparation. Furthermore, it is worth noting that the concept of middle school may be more clearly defined and prevalent in certain regions of the world, while in other countries, it may have less established foundations. For instance, as highlighted by Dickson et al. (2018), the United States stands out as a leader in MYP implementation, which is unsurprising given the prevalence of middle schools there. Consequently, the relative abundance of middle schools in the United States, where MYP implementation is more prevalent, might also contribute to the disparity in research about the MYP across the world. Importantly, although the MYP has claimed an advantage in fostering critical thinking, no research has evaluated the MYP's impact on students' critical thinking skills across varied contexts and relative to similar students enrolled in national curricula. There has been only a single study that has compared critical thinking across IB and non-IB samples and this was conducted among the IB DP (Cole et al., 2015), and an IB advantage was found among IB students. However, this study was

limited since its authors did not account for student and school characteristics. This was subsequently examined by Double et al. (2023) who, after accounting for several confounding variables, identified an advantage for DP students in their critical thinking when compared with their non-DP peers. To date, no such study has been conducted in the MYP. Therefore, in our study, we respond to Dickson et al.'s (2018, p. 254) call for 'comparisons of IB and non-IB students and/or schools with adequate controls'. Consequently, through the following questions, our research examined the critical thinking skills of MYP students compared with peers in national secondary education programmes (non-MYP[2]) in Australia, England and Norway:

1. Do MYP students differ from their matched non-MYP peers in their critical thinking skill?
2. What is the relation between MYP participation and critical thinking skill as a function of grade and country?

By addressing these questions, we hope to shed light on the effectiveness of the MYP in developing critical thinking skills and provide insights into how best to teach this vital skill in a variety of contexts.

MATERIALS AND METHODS

Participants

In total, 870 (386 MYP and 484 non-MYP) students participated from 21 schools across Australia, England and Norway. We adopted a simple random sampling technique, and, in each school, every student had an equal chance of participating in the study. We identified a comprehensive list of all MYP schools. This list served as our sampling frame, ensuring that all eligible schools were considered. We initially assigned a unique identifier to each school offering MYP 4 and MYP 5 and employed a random number generator to select specific schools for initial contact. Subsequently, we repeated this process, considering the feedback received from schools regarding their interest and willingness to participate in the study. Students were in either MYP 4 or MYP 5 (aged 14 to 16 years), which corresponds with Grades 9 (44.2%) and 10 (55.8%), respectively, in the national programmes. A breakdown of students by country and programme is provided in Table 2.

Thirty-one schools were initially eliminated as they did not offer the MYP 4 (equivalent to Grade 9) and MYP 5 (equivalent to Grade 10); there were 58 remaining MYP schools across the three countries that were eligible for participation. Schools that had an authorisation date up until 2017 (or close to if beyond 2017) were also prioritised, given that this was conceptualised as sufficient time for the MYP to be established within the school. However, given the difficulties with recruitment because of the pandemic, there was leniency in the authorisation year for three schools (one in Australia with an authorisation year of 2018 and two in Norway with authorisation years in 2018 and 2019) (see Table 4).

TABLE 2 Student sample size as a function of country and IB status.

	Non-IB	IB
Australia	249 (29%)	195 (22%)
England	123 (14%)	100 (12%)
Norway	112 (13%)	91 (10%)

All three schools had other IB programmes at varying levels (e.g., PYP or DP, or both) that were authorised well before the MYP. Therefore, these schools were quite familiar with the overarching IB frameworks and methodology. In addition, school selections in Australia presented a unique case in that public (state) schools were not prioritised given the complex and lengthy process of gaining ethical approval through the local Education Directorate.

The nature of the present study requires a comparison between MYP and non-MYP schools. However, it is not commonplace for schools that offer the MYP curriculum to also offer a separate national (state) curriculum during the middle years within that same school to a separate group of students. Based on the inclusion/exclusion criteria, eight MYP schools from each of the three countries were initially selected and were then approached to determine their interest in participating. To widen participation, and to account for potential attrition, especially due to the uncertainty of the pandemic, we also over-recruited in each country, particularly in Australia and Norway, where there were high attrition rates or unavailability to devote resources to participate in the study. In total, headmasters/principals from MYP schools in Australia ($n = 29$), England ($n = 8$) and Norway ($n = 17$), were invited to participate in the study. In total, and across the three countries, invitations were sent to 54 MYP schools (approximately 90% response rate).

Based on which schools indicated interest and provided consent to participate, a list of potential comparison schools that broadly matched the participating IB schools' socioeconomic background were compiled. Then, headmasters/principals from non-MYP schools in Australia ($n = 91$), England ($n = 39$) and Norway ($n = 92$) were invited to participate in the study. This included a mixture of public and private schools, except for Australia, where only private schools were contacted, given the length of time that it would take to obtain ethical clearance for public (state) schools. In total, and across the three countries, invitations were sent to 222 non-MYP schools (approximately 65% response rate).

Materials

Cornell critical thinking test (CCT)

A 52-item multiple choice critical thinking test covering 'induction, deduction, evaluation, observation, (statement) credibility, assumption identification, and meaning' (Ennis et al., 2005, p. 2). The test takes approximately 50 min to administer and is divided into subsections with moderate to high reliability (0.67 to 0.90) and split half reliability (0.55 to 0.76). Several studies have also confirmed the construct validity of the test (e.g., Behar-Horenstein & Niu, 2011; Ennis, 1993; French et al., 2012; Michael et al., 1980). The Level Z of the CCT was administered in the current study, which is recommend for high school students (Walker & Kettler, 2020). Table 3 provides a breakdown of the age ranges that are typical for high school students across the different countries represented in the current study. Although the United States is not a part of the sample for the current study, these US-based age ranges are presented as a baseline since this is where the CCT was normed. In particular, the option to administer Level Z considered the time of year in which the assessment was administered. More specifically, students were towards the end of their academic year, which meant that they were transitioning to the next age (e.g., 14-year-olds were becoming 15-year-olds). This was crucial to consider, as it influenced our choice of the appropriate CCT level in the current study, ensuring that we administered the level that most closely aligned with the students' developmental stage. There were other reasons motivating our option to administer Level Z. Firstly, the reduction in the number of items from 71 items (Level X) to 52 items (Level Z) was important for online administration,

TABLE 3 Typical student age levels (in years) according to grade across participating countries.

Country	Grade 9	Grade 10	Grade 11	Grade 12
USA	14/15	15/16	16/17	17/18
Australia	13/14	14/15	15/16	16/17
England	13/14	14/15	15/16	16/17
Norway	14/15	15/16	16/17	17/18

Note: In the current study, data was not collected from students in the USA. This information is only included as baseline information because this is where the CTT-Z was normed. This information is useful as it provides insight into the extent to which the Level Z of the CCT can be validly and reliably used in the target population in the current study (other three countries listed in the table above).

TABLE 4 Overview of schools, task administration details and MYP authorisation year (when applicable).

School location and type	Location	Delivery of task	MYP authorisation year
1. Australia (Private)	Onsite	Same day	2018
2. Australia (Private)	Onsite	Same day	2017
3. Norway (Public)	Onsite	Same day	2008
4. Norway (Public)	Onsite	Same day	2018
5. Australia (Private)	Onsite	Across 2 days	N/A
6. Australia (Private)	Onsite	Same day	N/A
7. Australia (Private)	Onsite	Same day	N/A
8. Australia (Private)	Onsite	Same day	N/A
9. England (Private)	Remote	Same day	1995
10. England (Private)	Onsite	Same day	2004
11. England (Private)	Onsite	Same day	2000
12. England (Public)	Onsite	Same day	N/A
13. England (Public)	Onsite	Same day	N/A
14. Norway (Private)	Onsite	Across 2 days	2008
15. Norway (Public)	Onsite	Across 2 days	N/A
16. Norway (Public)	Onsite	Across 2 days	2019
17. Norway (Public)	Onsite	Same day	N/A
18. Norway (Public)	Onsite	Same day	N/A
19. England (Public)	Onsite	Same day	N/A
20. England (Private)	Onsite	Same day	2015
21. England (Public)	Onsite	Same day	N/A

thereby streamlining the assessment process. This streamlining was particularly pertinent in the context of remote testing during the COVID-19 pandemic. Secondly, the use of Level Z was previously successful with populations in both English and non-English speaking contexts, which holds relevance for the Norwegian sample in our current study. Importantly, the language of the test was age-appropriate, and any potentially ambiguous aspects were clarified within the CCT itself (i.e., students did not need expert background knowledge to engage with the test).

The CCT score served as the dependent variable in the current analyses. Scoring of the CCT is based on a right only method. This involves counting the number of correct answers given by the test-taker. Each correct response to a test item is assigned one

point. Additionally, the measures described below were incorporated as control variables to mitigate potential observed differences between students that might influence their critical thinking skill. For example, it is important to control for cognitive abilities and intelligence as prior research have established an association between these constructs and critical thinking (e.g., Halpern & Butler, 2018; Halpern & Dunn, 2021; Stanovich & West, 2008).

Socioeconomic status

Participants provided several indicators of socioeconomic status (SES), including parental education, parent occupation and number of books in their household (Heppt et al., 2022). Students from different socioeconomic backgrounds may experience varying levels of critical thinking or access to developing this skill. It is essential to account for this factor to ensure that the study's findings are not confounded by socioeconomic disparities. One measure that has been often linked to SES is the number of books in the household (Heppt et al., 2022); this has also been used for international large-scale assessments such as the Progress in International Reading Literacy Study (PIRLS) (Mullis et al., 2019). We adopted this approach, and, in the current study, students provided information about the number of books in the home. To provide more context for students, and with previous findings that higher achieving students are more likely to better estimate the number of books in the household (e.g., Eriksson et al., 2022), we provided visualisations of books mapped against the respective response options: none or very few (0–10 books), enough to fill one shelf (11–25 books), enough to fill one bookcase (26–100 books), enough to fill two bookcases (101–200 books), enough to fill three or more bookcases (more than 200). Parents reported on their level of education according to the following options: prefer not to say, no I did not attend university, yes, I attended university and have an undergraduate degree or yes, I attended university and have a postgraduate degree. If applicable, the parent/guardian completing the form was also asked to select the best option that represented their partner's educational status. The reporting parent was also asked to provide details on their occupation as well as their partner's occupation. Due to the large amount of missing data for the parental education variable (approximately 75%) and the parental occupation variable (approximately 87% for the reporting parent and approximately 82% for details on their partner's occupation), they could not be used in the quantitative analyses and so the student reported number of books in their household was used as the only proxy for SES.

Big-5 personality assessment

Previous studies have linked personality to critical thinking (Acevedo & Hess, 2022; Clifford et al., 2004). For example, hierarchical multiple regression analyses identified a robust link between the personality dimension of openness to experience and critical thinking (Clifford et al., 2004). Therefore, to account for this potential influence of personality, a 50-item personality battery was administered using items drawn from the International Personality Item Pool (Goldberg et al., 2006). The inventory is based on the five-factor model of personality (Goldberg, 1993) and consists of ten items for each of the five personality dimensions: neuroticism, extraversion, conscientiousness, agreeableness and openness to experience. Participants rated each item on the extent to which it described them (extremely inaccurate to extremely accurate). The assessment has high reported internal reliability, with an average alpha value of 0.84 (Goldberg, 1993).

Cognitive reflection task (CRT)

Participants completed the three-item CRT. The task consisted of three questions taken directly from Frederick (2005) that have an intuitive yet incorrect response. The questions assess the participants' ability to deeply process problems and override intuitive thinking.

The international cognitive ability resource (ICAR-16)

Participants also completed the ICAR-16 (Dworak et al., 2021). The ICAR-16 is a 16-item broad intelligence measure of four aspects of intelligence: verbal reasoning, letter and number series, matrix reasoning and three-dimensional rotation. All items were in multiple choice format and the presentation of items was randomised.

Procedure

All participants were provided with information sheets and consent was gained from participants before engaging in data collection. To widen participation, the research team accepted invitations to deliver an overview of the study and what participation entailed to students. Given the difficulty in recruiting schools in Australia, a member of the research team also coordinated with the IB Schools Australasia team to advertise the research project in their newsletter. For the non-IB schools, a member of the research team reached out to headmasters and middle-years coordinators. In addition, due to school closures in response to the COVID-19 global pandemic, the entire study was remotely conducted. However, note that although the research team did not visit the physical location for task administration, some schools administered the task as an in-person session (onsite) in a quiet room on the school grounds (see Table 4).

Prior to commencing task administration, a research member of the team met online with the contact person from each school for 20 to 30 min to discuss the nature of participation and to agree on the logistics of gaining consent and administering the task. When a school decided on the specific day(s) and time(s) that students would engage with the tasks, this was communicated to a member of the research team so that they could be available at that time in case the school had any questions during task administration. Before this date, parents and students indicated their interest in participating in the study through the completion of a parent and student consent form, respectively. Given that parent/student responses were directly shared with the research team, to manage the consent process a member of the research team created individual, secure folders for each participating school. This folder contained three key documents: (1) consent status of students; (2) consent status of parents; and (3) key instructions about participating in the study, including the links to the student background questionnaire and critical thinking test. This folder was only accessible to the key contact persons from a specific participating school and the research team. Having access to this folder also enabled contact persons from each school to be aware of students who wanted to participate so that they would know the appropriate number of rooms to prepare before the date of task administration. This was particularly important, given that, in some instances, there were restrictions on the number of students that could be in a room at once.

Participants first completed the student consent, the questionnaires and cognitive task in the following order: demographic variables, socioeconomic indicator, personality, ICAR-16, then the CRT. All students subsequently completed the critical thinking test. All sessions, whether onsite or remote, were supervised by at least one teacher from each of

the participating schools and a research member from the research team was on standby, through Microsoft teams, on the day of task administration if any questions or technical difficulties emerged. Except for one school (that completed the task remotely), all students completed the tasks in a quiet room at their school. The school that completed the task remotely required all participating students to log into the school's secure online platform and students individually completed the task while being remotely supervised. Seventeen schools were able to accommodate student completion of both the student background questionnaire and critical thinking task on the same day. In the remaining four schools, students completed both tasks over a period of two days to accommodate timetabling requirements and supervision availability (see Table 4). In the case of the latter, students completed the consent form and the student background questionnaire on the first day (approximately 30 to 40 min for administration), followed by the critical thinking test on the second day (approximately 50 min). This sequencing was chosen to align with the time constraints associated with each task and the scheduling needs of these schools. Furthermore, items in the tasks for the student background questionnaire were presented in a random order for each student. For instance, the questions in the personality test were randomised to reduce order effects.

Data was subjected to exploratory multi-level regression analyses and propensity score matching using R, a language and environment for statistical computing (R Core Team, 2013). Propensity score matching was employed using a full matching procedure to balance the MYP and non-MYP groups concerning the background characteristics, including personality traits, SES, gender, and age. The rationale for adopting a propensity score matching approach was to simulate a randomised experimental scenario, in which the treatment group (i.e., students enrolled in the MYP) and the comparison group (i.e., students not enrolled in the MYP) exhibit similarity in relevant background characteristics (Stuart & Green, 2008). Propensity score matching is particularly valuable when random assignment to different groups is unfeasible, as was the case in our study. By implementing this matching strategy, we effectively aimed to reduce the influence of background characteristics, thus enabling more confident attributions of outcomes to the MYP curriculum, rather than any potential background differences between the groups. The overarching goal is to select a subgroup of students from an external comparator group (i.e., non-MYP students exposed to the national curriculum in each country) who share predefined characteristics (i.e., important covariates) with students in the study population exposed to the treatment (i.e., MYP students). Successful matching should result in covariates no longer predicting treatment group status. This approach emulates the randomisation process of experiments by matching students based on pre-defined criteria.

RESULTS

Descriptive analysis

The demographic composition of the sample as a function of IB participation is presented in Figure 1. Preliminary analyses suggested that the IB sample differs from the non-IB sample in terms of gender, $\chi^2 = 207.02$. Males were highly over-represented in the non-IB programme sample (69%) compared to the IB programme sample (20%). The distribution of MYP and non-MYP participants exhibited a generally similar pattern across all three countries, with an equal percentage (51%) of both MYP and non-MYP students in Australia. There was no difference in grade level as a function of IB participation, $\chi^2 = 1.77$; Year 9 students constituted 42% of the non-IB programme and 47% of the IB programme sample. A Wilcoxon rank sum test suggested that the number of books in the household differed between the IB and non-IB samples, $W = 73{,}457$, with IB students reporting a higher average number. In addition,

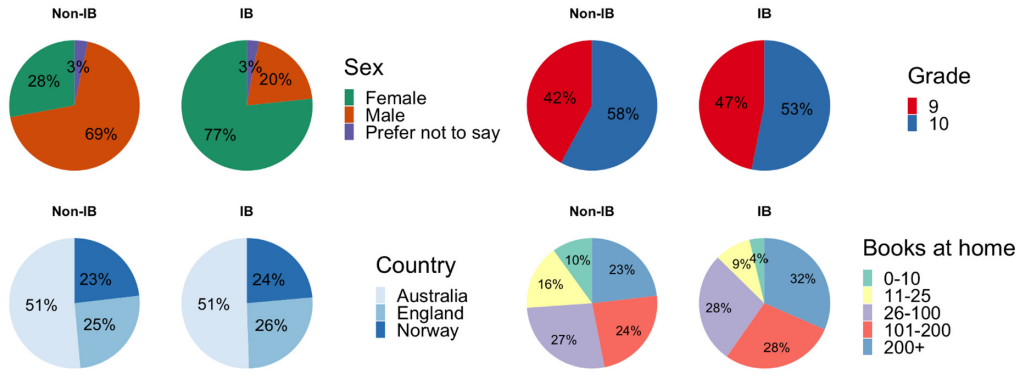

FIGURE 1 Demographic variables as a function of IB status.

Grade 9 ($M = 2.38$, $SD = 1.40$) and Grade 10 ($M = 2.53$; $SD = 1.54$)[3] MYP students have been enrolled in the IB for a similar number of years.

Differences in critical thinking

An initial analysis was performed to examine critical thinking differences between MYP and non-MYP students without accounting for confounding variables. Critical thinking was higher in the MYP ($M = 24.24$, $SD = 6.19$) compared to non-MYP students ($M = 21.12$, $SD = 6.54$), $d = 0.49$. Propensity score matching was then performed using a full matching procedure between the MYP and non-MYP groups in terms of several potential confounders (cognitive abilities, personality, SES, gender, and age) to obtain a set of matching weights. After applying a regression with these matching weights and using robust standard errors that account for clustering in the weighting classes, MYP participation was still a positive, moderate predictor of critical thinking, $\beta = 0.38$.

Exploratory multi-level regression analyses were performed to examine differences in critical thinking performance after accounting for covariates. The following variables were entered into the model: cognitive abilities (intelligence and cognitive reflection), personality variables (openness, conscientiousness, neuroticism, extraversion, and agreeableness), SES indicator (number of books at home), and demographic variables (age and sex). Age was only recorded in whole years rather than as a continuous measure. School was entered as a grouping factor. Importantly, for this and subsequent analyses, students with incomplete data had their missing responses handled using random forest imputation. This was deemed appropriate as very few variables had missing data, and those that did had 2% or fewer missing cases. Sample sizes for each model are noted where appropriate. The multi-level regression indicated that intelligence (ICAR 16), the highest two levels of cognitive reflection, and openness to experience were positive predictors of critical thinking, while extraversion was a negative predictor. The highest level of number of books at home was also a positive predictor of critical thinking. Crucially, after accounting for all covariates, IB status was a predictor of critical thinking, such that students in the IB programme had higher critical thinking. This effect was moderate in size, $\beta = 0.33$.

Finally, we performed a propensity score matching analysis. To determine relevant covariates for calculating a propensity score, we relied on both theoretical considerations and the empirical findings from a logistic regression with each of the covariates in the prior regression model but using IB participation as the criterion variable (see Table 5). Intelligence, openness to experience, conscientiousness, sex, and number of books at home were

TABLE 5 Results of the logistic regression analysis on IB participation without matching weights applied.

Predictors	Odds ratios	Std. beta
(Intercept)	0.27	1.21
Sex (male)	0.09	0.09
ICAR 16	1.13	1.58
CRT (1)	1.08	1.08
CRT (2)	0.97	0.97
CRT (3)	0.7	0.7
Openness	1	1.72
Conscientiousness	1	0.76
Extraversion	1	1.03
Agreeableness	1	1.04
Neuroticism	1	0.87
Age (14)	1.99	1.99
Age (15)	0.67	0.67
Age (16)	1.09	1.09
Age (17)	30.94	30.94
Number of books (11–25)	1.24	1.24
Number of books (26–100)	2.15	2.15
Number of books (101–200)	2.65	2.65
Number of books (200+)	2.75	2.75
Observations	870	
R^2	0.369	

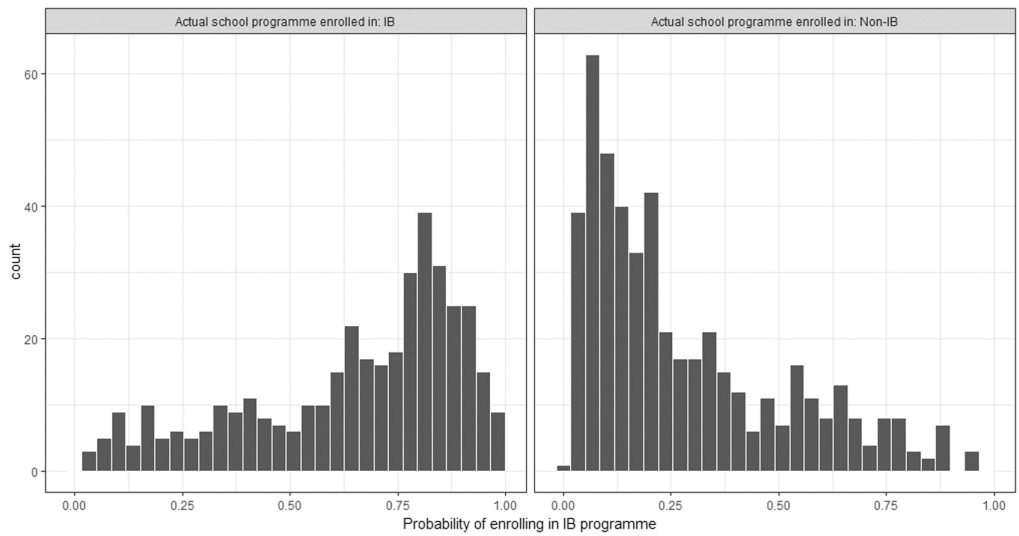

FIGURE 2 Distribution of propensity scores in the IB and non-IB samples.

predictors of IB enrolment, such that IB students tended to have higher intelligence scores, have higher openness to experience and conscientiousness, were more likely to be female, and they tended to have more books in their household. See Figure 2 for the distribution of propensity scores based on the model presented in Table 5.

We then calculated a propensity score based on the covariates in Table 5. The motivation for using a propensity score matching approach was to approximate a randomised experiment scenario where the treatment (i.e., enrolled in MYP) and comparison (i.e., not enrolled in MYP) groups are similar on all relevant background characteristics (Stuart & Green, 2008). Specifically, an optimal full matching procedure was used as this approach has been argued to be the best matching technique (Hansen, 2004). This approach uses all available people in the sample and groups them into matched sets (i.e., subclasses) based on their similarity across the covariates used for matching and each of the subclasses contains at least one 'treated' individual (IB) and one 'comparison' individual (non-IB). Weights are then assigned to each of the subclasses to 'equalise' the covariates across the treated and comparison groups, with more similar subclasses weighted more than less similar ones. If the matching procedure is successful, the covariates should no longer predict treatment group status.

In our case, the matching success was evaluated both in terms of the average differences in the covariates between the treatment and comparison groups before and after the matching procedure was applied, as well as by re-running the logistic regression presented in Table 5. The full matching procedure was highly successful and was able to reduce all covariate differences to an acceptable level, and this was particularly marked for the average differences in sex, intelligence, and openness to experience in the unmatched IB and non-IB samples, which was important given the findings of the initial logistic regression on IB participation. Table 6 presents the re-run logistic regression with the matching weights applied, showing that none of the covariates now predict MYP participation status. To evaluate the grade and country-level comparisons, the same matching approach was carried out for

TABLE 6 Results of the logistic regression analysis on IB participation with matching weights applied.

Predictors	Odds ratios	Std. beta
(Intercept)	0.9	1.27
Sex (male)	1.04	1.04
ICAR 16	1.01	1.04
CRT (1)	0.89	0.89
CRT (2)	0.97	0.97
CRT (3)	1.04	1.04
Openness	1	1.03
Conscientiousness	1	1.16
Extraversion	1	0.92
Agreeableness	1	0.97
Neuroticism	1	1.03
Age (14)	1	1
Age (15)	0.68	0.68
Age (16)	0.73	0.73
Age (17)	3.03	3.03
Number of books (11–25)	0.86	0.86
Number of books (26–100)	0.77	0.77
Number of books (101–200)	0.88	0.88
Number of books (200+)	0.89	0.89
Observations	870	
R^2	0.013	

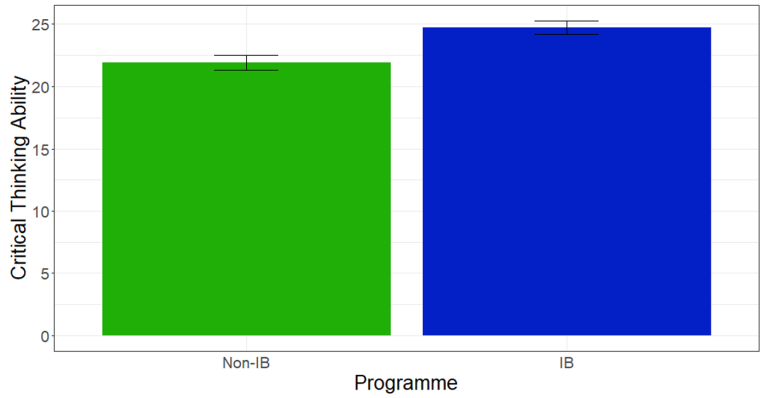

FIGURE 3 Weighted average of the critical thinking scores (plus and minus the weighted standard error of the average) as a function of IB MYP participation.

these subsamples. Given fewer cases with which to match in these subsamples, it is unsurprising that the matching tended to be less effective for them, although it was generally still of high quality. The matching for the Norwegian subsample was the exception to this, and so their country-level findings should be interpreted with a degree of caution (see supplementary materials, Figures S1-S5, for details on the matching procedure for the subsamples).

The role of IB participation status on critical thinking was then investigated as per the recommendation of Ho et al. (2007) by applying a regression analysis on the CCT scores including IB status and the matching covariates as predictors with the matching weights applied. After matching the IB and non-IB samples on these various covariates, participation in the MYP was identified as a positive predictor of critical thinking ability (see Figure 3) and this effect was moderate in size, $\beta = 0.38$.

We further examined whether the differences between MYP and non-MYP students differed as a function of grade level and country through a series of additional weighted regressions.

Grade level comparisons

We examined whether the differences between IB and non-IB students differed as a function of grade level and the findings are illustrated in Figure 4. For Grade 9 students, IB status was again a positive predictor of critical thinking, and this effect was moderate in size, $\beta = 0.37$. Similarly, for Grade 10 students, IB status was a positive predictor of critical thinking, and this effect was moderate in size, $\beta = 0.33$. As illustrated in Figure 4, both the overall performance and difference between IB and non-IB participants were very similar between the grades.

Country-level comparisons

Finally, we examined whether the differences between IB and non-IB students differed as a function of country. In Australia, IB status was a positive predictor of critical thinking, and this effect was moderate in size, $\beta = 0.30$. Similarly, in England, IB status was a positive predictor of critical thinking and this effect was moderate to large in size, $\beta = 0.62$. In Norway, IB status negatively predicted the critical thinking score, $\beta = -0.06$. As Figure 5 illustrates, in general, the non-IB students' average performance was very similar across the three countries, but

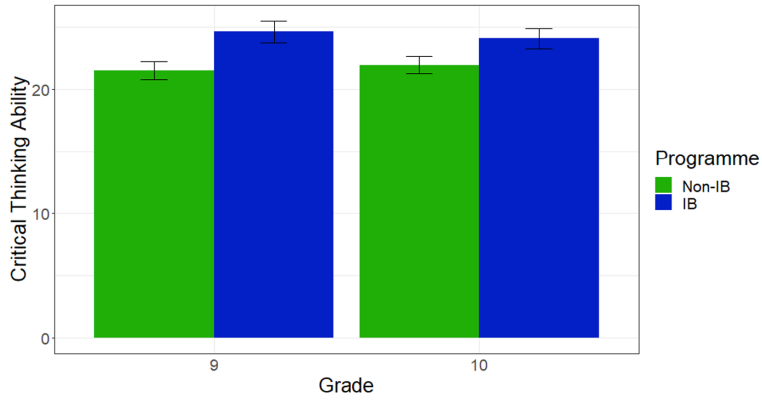

FIGURE 4 Weighted average critical thinking score as a function of grade and IB MYP participation. Grade 9: non-IB ($M=21.51$; $SD=6.36$), IB ($M=24.63$; $SD=6.82$); Grade 10: non-IB ($M=21.97$; $SD=6.95$), IB ($M=24.08$; $SD=6.18$); *M*, mean; *SD*, standard deviation.

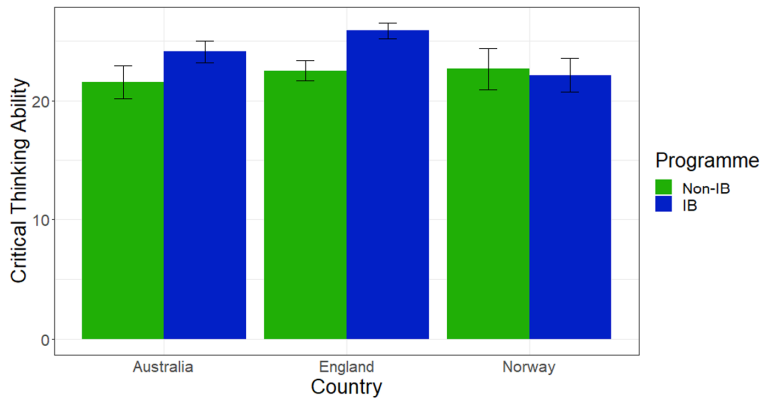

FIGURE 5 Weighted average critical thinking score as a function of country and IB MYP participation. Australia: non-IB ($M=21.54$; $SD=7.72$), IB ($M=24.08$; $SD=5.47$); England: non-IB ($M=22.51$; $SD=5.32$), IB ($M=25.85$; $SD=4.54$); Norway: non-IB ($M=22.64$; $SD=7.74$), IB ($M=22.12$; $SD=8.44$); *M*, mean; *SD*, standard deviation.

for MYP students, it was noticeably higher in England and particularly compared to the Norwegian matched sample average. Similarly, the difference between the IB and non-IB students' performance was noticeably larger for the English matched groups, as was also reflected in the substantially larger effect size for this difference, compared to the Australian and particularly the Norwegian matched groups.

DISCUSSION

Over the past years, critical thinking has emerged as a fundamental objective in education, prompting educators and scholars to tackle the difficulties involved in effectively teaching this essential skill. It is not sufficient to incorporate critical thinking into the curriculum; instead, creating the appropriate conditions for students to learn and develop this competency is imperative. In the present study, we compared the critical thinking skills of MYP and non-MYP, national curricula students. The current study demonstrated that MYP students

had higher levels of critical thinking skill and with a moderate role of IB participation. After accounting for important pre-existing differences (gender, personality and general cognitive abilities), we observed that there was still an advantage in critical thinking for the IB sample. This finding remained consistent across grade level (Grade 9 and Grade 10) and across English and Australian students. However, there was no meaningful difference for the Norwegian country-level comparison.

The case of connected explicit instruction and assessment in critical thinking development

There has been a long-standing debate about the most effective way of teaching critical thinking, which stems primarily from disagreements about its theoretical conceptualisations (Alazzi, 2008; Chirgwin & Huijser, 2015; Munkebye & Gericke, 2022; Vieira et al., 2011). That there was a MYP advantage in critical thinking development might be associated with the presence and nature of explicit instruction in the MYP compared with national curriculum approaches. This is a noteworthy observation because previous meta-analyses have underscored explicit instruction as a fundamental catalyst for the development of critical thinking. A pivotal aspect of this involves the clear and explicit definition of what constitutes critical thinking (Abrami et al., 2008, 2015). MYP teachers and coordinators in England, some of whom have experience with both national and MYP curricula, have shed light on the distinctions in critical thinking instruction within MYP and non-MYP contexts (McGrane et al., 2022). Interestingly, they noted that the nature of the content does not significantly differ between MYP and non-MYP schools (McGrane et al., 2022). Instead, it was emphasised that the MYP offers a more explicit and effective approach to providing students with opportunities for active thinking and reflection. MYP teachers have further advanced two benefits of employing explicit assessment language to foster critical thinking development: (1) these criteria serve to emphasise the significance of critical thinking to students; and (2) it facilitates a more transparent understanding of expectations among students and between students and teachers (McGrane et al., 2022; Wright et al., 2016).

The effectiveness of explicit instruction, especially in relation to critical thinking, may be attributed in part to Sweller et al.'s (2019) cognitive load theory. Sweller et al. (2019) suggest that explicit instruction and feedback are effective because humans have evolved to learn directly from others through the borrowing and organising principle, which reduces the load on working memory. This is particularly helpful for students who may struggle with learning, such as those with learning difficulties or from linguistically diverse backgrounds, when presented with complex and new information, like the critical thinking strategies of when, why, and how to use them (Chen et al., 2018; Siregar, 2021). When information is received and stored in long-term memory, the constraints of working memory are diminished, allowing an effective application of feedback students. Sweller et al. (2019) argue that by designing instructional experiences that correspond with this cognitive process, the acquisition of skills can be improved. The MYP strengthens the development of critical thinking by providing students and teachers with a common and clear language to engage in the borrowing and organising process through its criterion-referenced descriptors and strands, which are utilised in project-based assessments such as the MYP personal project (IBO, 2012, 2022). This aligns with previous studies which have shown that teachers and school leaders have favourable views regarding the IB assessment frameworks and practices and have acknowledged that this enhances the delivery of explicit instruction in fostering critical thinking (Abrami et al., 2008, 2015; Dickson et al., 2018; Marin & Halpern, 2011). For example, Visser's (2010) research on IB assessment in multiple countries including Australia, the Netherlands, Canada and the USA, demonstrated that many MYP coordinators and

teachers considered assessment to be the most valuable aspect of the MYP. Moreover, Wright et al. (2016), across 54 countries, found that 70% of the IB coordinators regarded the availability of various explicit assessment instruments in the MYP as highly valuable.

Notwithstanding the importance of the explicit approach, MYP teachers and students alike report that the IB's integration of explicit instruction within a connected learning framework fosters critical thinking (McGrane et al., 2022). This was frequently attributed to the emphasis on interdisciplinary learning and the establishment of cross-curricular connections (Hamza-Lup & Sopin, 2008; Kaklanis et al., 2015; McGrane et al., 2022). For example, students from England, Norway and Australia identified that explicit instruction within a connected framework permitted them to transfer their learning of important practices from key units such as Individuals and Societies (INS), including source analysis, to enable more critical analysis and synthesis across other subject areas (McGrane et al., 2022). Similarly, 79% of MYP coordinators in an extensive cross-national study, which included Australia, reported an appreciation for the MYP's prioritisation of interdisciplinary teaching (Wright et al., 2016). More broadly, MYP teachers from the three countries under consideration in this study also reported that connected learning was leveraged in a way that permitted them to better manage time on task: they recognised that critical thinking strategies learnt in INS, for example, permitted students to more quickly engage and complete tasks requiring evaluation and analysis (McGrane et al., 2022).

This connected learning framework is driven by the MYP's approaches to learning (ATL's) (IBO, 2014a). More relatedly, within the ATL's thinking skills category, critical thinking is recognised as an essential skill that students must acquire to excel as learners. This is underpinned by inquiry-based learning, which is widely recognised as an effective approach for fostering critical thinking (Evangelisto, 2021; Fuad et al., 2017; Soneral & Wyse, 2017; Wahyudi et al., 2019; Zandvakili et al., 2019). Inquiry-based approaches employ a student-centred process which helps them to develop their own conceptual knowledge and frameworks (Anderson et al., 2013). Within these approaches, there is an emphasis on authentic problem-solving within a connected and collaborative learning environment (Ku et al., 2014). This is indeed a hallmark of the MYP, which is known for its student-centred approach that offers opportunities for students to develop their own sense of agency and independence in the process of becoming critical thinkers (Sperandio, 2010; Taylor & Porath, 2006; Waterbury, 2018). Alongside this, teachers also engage in their own process of connected learning through the inquiry-based framework, which is a considerable shift for most educators who have previously taught in schools that are usually teacher-centred (Grewal, 2001; Hill, 2001; Sperandio, 2010). For example, in Norway, MYP teachers have reported that the interdisciplinary units offer an opportunity for students to extend their metacognition and abstract thinking process (Termaat, 2023). Therefore, the MYP experience for teachers is as much an inquiry-based, self-directed process as it is for students. While most classrooms and curricula include some of the above-mentioned strategies to foster student critical thinking (as evident in the introductory section discussing the national curricula in Australia, England and Norway), the MYP has perhaps claimed an advantage in fostering critical thinking through its systematic and strict prioritisation of explicit instruction through a connected learning framework underpinned by inquiry-based learning (IBO, 2014a).

Interestingly, the MYP's connected, explicit instructional approach to critical thinking instruction, as described above, diverges from the methods found in the national curricula of the countries under study (except for Norway, where we did not identify a MYP advantage). Intriguingly, our research reveals a significant IB advantage, primarily observed among English-speaking students, in contrast to their peers in Australia and Norway. But why might this be the case? One key reason might be related to the nature of explicit instruction in England under the national curriculum. Unlike Australia and Norway, England's national curriculum and teacher training neither explicitly defines nor positions critical thinking as a core

educational goal (Belas, 2019; DfE, 2013a). Admittedly, an examination of the DfE subject curriculum guidance (DfE, 2013b, 2015) demonstrates that each subject aims to develop critical thinking. Since most students in England will prepare to sit the General Certificate of Secondary Education (GCSE) at the age of 16 through the national curricula route, they will be offered an opportunity to develop their critical thinking (Ofqual, 2017). However, Sizmur and Cunningham (2013) highlight that in England, although there is evidence in the GCSE (aligns with national curricula) documentation of the importance of progression student higher order thinking skills, there was less explicit expectations for the development of these skills when compared with international curricula such as the MYP. Though the impact of MYP participation was most noticeable in England, there were also clear advantages for the IB in Australia. This could be related to the fact that within the Australian national curriculum landscape, critical thinking is delivered in a compartmentalised, implicit fashion (NSW Board of Studies, Teaching and Educational Standards, 2014; Queensland Curriculum and Assessment Authority, 2015; VCAA, 2015). This implicit approach is at odds with previous recommendations of best practices in critical thinking instruction, which could also explain the absence of a national curricula advantage in Australia in the current study.

These findings rekindle the persistent debate about whether critical thinking is most effectively taught explicitly—the 'generalist view' (infusion approach)—or implicitly (immersion approach)—the 'specifist view' (Davies, 2006, p. 1; Ennis, 1989). For example, in Australia, after a long history of failing to maintain critical thinking instruction as an explicit, cross-curricular endeavour (Gilbert, 2019; also see Watt, 2006 for the Tasmanian 'Essential Learnings'), and despite calls from the Gonski report, which recommended that these skills be placed 'at the core of our [national] curriculum and teaching practice' (Gonski et al., 2018, p. 38), it was only in recent years that the Australian national curriculum changed its recommendation of teaching critical thinking in a subject-specific fashion (specifist view) to one that views the skill as a general curriculum capacity (generalist view) (Cole et al., 2015). Indeed, Australian schools that have discontinued the MYP noted that curricular changes are ensuring that the Australian national curriculum now have 'a lot IB-ness to it' as the '(Australian Curriculum) really drew a lot of its ideas from IB' (Dickson et al., 2021, p. 18). Nevertheless, while critical thinking in Australia is viewed as a valued educational goal, contradictory views that seem to side with a specifist view have remained in circulation, with an official review of the Australian national curriculum recommending that 'critical and creative thinking … should be embedded only in those subjects and areas of learning where relevant and where they can be dealt with in a comprehensive and detailed fashion' (Donnelly & Wiltshire, 2014, p. 248). This raised more questions than answers, since the review's authors neither specified how to position the teaching of critical thinking within the Australian national curriculum nor did they provide convincing evidence-based research to support their recommendation. It is therefore unsurprising that although Australian secondary school teachers endorse the relevance of critical thinking, much like British national curriculum teachers, they do not consistently plan for the implementation of explicitly teaching this skill (Skourdoumbis, 2016). Moreover, evidence suggests that students may require more support to acquire critical thinking skills (Abrami et al., 2015), and many national curriculum educators continue to face challenges in integrating effective critical thinking instruction into their curricula due to a lack of expertise (Ab Kadir, 2018; Erikson & Erikson, 2019; Paul & Elder, 2019).

Intriguingly, the present findings further raise the question of whether integrating both a generalist (infusion) and specifist (immersion) pedagogy could hold the key to fostering critical thinking most optimally. Note that both the generalist and specifist view advocate for embedding the teaching of critical thinking across different subjects. This contrasts with the general approach (variation of generalist approach) which advocates for the delivery of explicit critical thinking instruction as a stand-alone course that is independent from a specific subject (Ennis, 1991; Gann, 2013). In a prominent meta-analysis, Abrami et al. (2008)

analysed 117 studies that comprised 20,698 participants and arrived at the conclusion that critical thinking skills were most optimally developed through explicit rather than implicit teaching methods, which treats critical thinking as an outcome (rather than the subject) of instruction. This finding has been consistently supported by research across the past 15 years across different contexts and employing different methodological designs, including randomised controlled trials (Arı, 2020; Arısoy & Aybek, 2021; Cole et al., 2015; Marin & Halpern, 2011; Smith et al., 2018). However, that Orhan and Çeviker Ay (2022) identified general, immersion and mixed [combined] (general and immersion) approaches significantly improved secondary-level students' critical thinking skills, all with large effect sizes, is interesting. Although, note that the magnitude of the effect favoured a general, followed by the mixed and then the immersion approach. While these results seem to favour the explicit approach (albeit via a standalone approach), they also indicate that combined approaches, such as those implemented in the MYP curriculum and Norwegian national curriculum (we examine the nuances of the Norwegian findings later in the discussion) seem to also be effective in fostering critical thinking. By contrast, the traditional, content-based curricula, which dominate in Australia and England, might position critical thinking instruction at the periphery, thus creating minimal opportunities for explicit instruction of this skill, especially if it is not positioned as accountable outcome (Cole et al., 2015).

Upholding explicit instruction: Issues of accountability and fidelity in curricula implementation

Beyond establishing a shared operational language for explicit instruction aimed at enhancing critical thinking, MYP educators observed that incorporating this explicit language into assessments helped cultivate a culture of accountability and fidelity in the MYP implementation (McGrane et al., 2022). In previous interviews, MYP teachers from all the schools included in the present study consistently emphasised the importance of explicitly communicating to students the various assessment strands and levels related to critical thinking. Notably, when it comes to fostering critical thinking, many of these teachers consistently referred students to Criteria A and D of the MYP assessment framework, which focuses on analysing and evaluating, two core components of critical thinking skills (Ennis et al., 2005; IBO, 2014a). This approach stands in contrast to the level of accountability typically observed in national curriculum strategies aimed at cultivating critical thinking. For example, in England, the national curriculum sits against the backdrop of England's history of introducing and removing courses specifically designed for critical thinking development. More specifically, in 2011, England's Oxford, Cambridge and RSA Examinations (OCR) exam board published a stand-alone module (Level 2 Thinking and Reasoning Skills) for secondary-school students (OCR, 2011), which was later abandoned in 2018 by the DfE because of government reforms (OCR, 2016). Although this may suggest a move to align with best practices (as indicated in our previous discussion on the tensions and consequences of teaching critical thinking through a siloed approach), this exemplifies an under-prioritisation of critical thinking in England's national curricular accountability goals, part of which may be attributed to their difficulty in conceptualising and, therefore, assessing critical thinking (Ventista, 2018). Moreover, that the British Council (2016) continued to implement standalone initiatives for critical thinking (e.g., Connecting Classrooms Critical Thinking and Problem-Solving programme) is testament to a lack of a unified understanding and accountability structure for critical thinking development across the British Government. Bhatt's (2022) findings support this view such that secondary school teachers in England perceived that greater accountability to develop critical thinking would motivate teachers to understand the nature of critical thinking, including its composition and effective pedagogical approaches.

However, this might not be an ideal solution, given that accountability measures, if not properly implemented or if based on misguided conceptualisations, may lead to didactic and narrow pedagogical approaches, which may have negative consequences for critical thinking development (Alwehaibi, 2012).

The issue of accountability and fidelity is further important when considering the absence of an IB advantage in Norway in the current study. It is firstly important to acknowledge that this could have been driven by the less effective matching for the Norwegian sample which we discuss later in our limitations. However, other factors related to the status of accountability in critical thinking instruction within the Norwegian national curriculum and at the MYP level may be at play. In relation to the national curriculum, the Norwegian education landscape differs from the previously highlighted contradictory messages within Australia and England: there is a more consistent message and a move towards greater accountability for the development of critical thinking (Norwegian Directorate for Education and Training, 2020). Our finding of no advantage for MYP Norwegian students in the present study likely reflects this cohesive integration of critical thinking within Norwegian national curricula (Scheie et al., 2022), which contrasts with the UK and Australia where there are not unified national approaches to critical thinking instruction, as discussed above. A key component of Norway's unified approach to fostering critical thinking is through a knowledge-based national curriculum (Ott, 2019; Scheie et al., 2022). More specifically, Norway recently underwent curriculum changes, with a view towards strengthening students' core background knowledge and narrowing its content to facilitate deeper learning (Norwegian Directorate for Education and Training, 2020; Norwegian Ministry of Education, 2017). This approach to strengthening background knowledge has been previously linked to better critical thinking outcomes (Ennis, 1989; Willingham, 2008), with the understanding that critical thinking skills cannot be achieved unless there is sufficient subject knowledge (Thomas & Lok, 2015; Willingham, 2008). Additionally, Vardi (2015) highlights the mutually reinforcing cycle of domain knowledge and critical thinking within a discipline. Consequently, developing critical thinking skills within a discipline is highly advantageous for both critical thinking and content knowledge development, since it enables individuals to establish meaningful connections between different parts of the subject matter, thereby leading to a deeper understanding of a topic. Moreover, although Norway prioritises the teaching and assessment of critical thinking through a generalist approach, the design of the curriculum offers students an opportunity to transfer their thinking across different situations and disciplines (Norwegian Directorate for Education and Training, 2020).

This process is also advocated within the broader MYP framework, and by extension within the MYP in Norway (IBO, 2012, 2014a, 2022), although, in the MYP, rather than referring to a knowledge-driven curriculum, they mention a concept-driven curriculum that is distinguished by key (overarching) concepts and related (subject-specific) concepts (IBO, 2014a). Despite the difference in name, both curricula present an explicit teaching of critical thinking and strengthening of knowledge or concepts (IBO, 2014a; Scheie et al., 2022). Importantly, both the MYP curriculum and the Norwegian national curriculum place a strong emphasis on addressing multi-categorical issues with ethical implications. Within the MYP IB curriculum, key concepts such as perspective, identity and communities are explored, along with related concepts such as empathy, bias, poverty and equity (IBO, 2012, 2014a, 2022). Similarly, the Norwegian national curriculum stresses the importance of ethical awareness, prompting students to consider the ethical implications of their decisions and actions (Norwegian Directorate for Education and Training, 2020). This knowledge-based approach facilitates the creation of learning environments that enable students to develop their ability to think more broadly. Encouraging students to contextualise academic ideas and concepts within a broader framework permits the establishment of meaningful connections between the dissemination of knowledge in the classroom

and real-world applications of academic concepts (Velez & Power, 2020). Furthermore, when these issues are approached through a dialogue-reflection-action process, students are motivated to think critically and apply their thinking in practical situations (Davies & Barnett, 2015). Interestingly, central to both the Norwegian and MYP curricula is the explicit integration of the learning and transfer across subject areas, which seemingly amalgamates both generalist and specifist approaches to teaching critical thinking. Indeed, previous scholars (e.g., Ennis, 2013; Tiruneh et al., 2017) have supported this combined approach to fostering critical thinking with some scholars opposing the rigid dichotomy between the generalist and specifist view and have suggested that the debate between these two schools of thought involves a 'fallacy of the false alternatives' (e.g., Chaffee, 1994; Davies, 2006, p. 2). This view aligns with those from cognitive science research suggesting that the learning of complex processes such as critical thinking benefits greatly from the initial explicit, direct instruction, such as those proffered by the generalist view (Lai, 2011). However, this is purely speculative and further qualitative research is required to explore the specific approaches that are being implemented in the current sample.

Challenges in maintaining fidelity in MYP implementation

If the MYP is clear on its approach to critical thinking development, we must return to addressing the question of why an IB advantage did not manifest in Norway. This could perhaps be related to the non-prescriptive nature of the MYP, where schools have the autonomy to tailor their objectives, content, and assessment strategies to their specific educational context (Sperandio, 2010). This openness can lead to the integration of local (national) curriculum components in a manner that may not align with the fidelity of MYP implementation. Previous interviews have revealed that this flexibility poses significant challenges for Norwegian MYP teachers and coordinators (McGrane et al., 2022). Admittedly, this experience is not unique to Norway; MYP educators in Australia (Perry et al., 2018) and the UK (Sizmur & Cunningham, 2013) have reported higher workloads as they grapple with the dual adoption of MYP and national curricula, including the complex task of managing two distinct assessment frameworks. However, what distinguishes the Norwegian context is the stringent legal mandate to integrate the Norwegian local curriculum into the MYP framework. Given the legal mandate and requirement for delivering Norwegian examinations, the emphasis often falls on prioritising the preparation for these examinations (McGrane et al., 2022). This requirement therefore has the potential to curtail the flexibility of MYP schools in Norway when contrasted with those solely following the MYP curriculum or enjoying the flexibility to do so, as observed in Australia and England (Dickson et al., 2020, 2021; McGrane et al., 2022).

MYP teachers and coordinators in Norway have further indicated that due to these pressures and the emphasis on the local national curriculum, they often extend specific subject areas beyond the recommended IB guidelines. This practice, however, can lead to other MYP subjects receiving less attention and potentially interfere with the requirement of at least 50 h of instruction for each subject group, as stipulated by the IB (IBO, 2022). For example, some Norwegian MYP teachers have reported that they perceive the process-oriented and documentation-heavy nature of the MYP may reduce students' motivation to engage in critical thinking, particularly in subjects like Design, where students are required to document step-by-step decisions before commencing the creative process. This requirement is less strict in other subjects, such as Languages and Literature (McGrane et al., 2022). Therefore, despite ongoing professional development and clear IB guidelines, some teachers may not consistently implement the MYP methodology due to time-related or legal-related constraints (Pendergast et al., 2014). Instead, Norwegian MYP teachers

sometimes opt to streamline the depth of student engagement in these activities, deviating from previous research recommendations that emphasise explicit involvement in the critical thinking process (Abrami et al., 2008, 2015). Future research should delve into understanding this concern and propose practical strategies for documentation to ensure that teachers implement the MYP with high fidelity (Bishop et al., 2014; Durlak, 2016; Hansen et al., 2013). In addition, future research may consider how the cultural background of teachers might influence the extent to which the MYP is implemented with fidelity.

Limitations and future directions

As with many complex studies that span multiple countries, there are limitations that should be addressed in future research. We used two measures to gather information about students' social, cultural and economic backgrounds: asking students how many books were in their household and asking their parents about their education levels. However, the use of the number of books indicator is concerning in this study because previous research has shown that there may be a gender bias in responses, with females reporting higher numbers than males of similar backgrounds (Engzell, 2021). We also attempted to collect data on parental education levels, which is a more accurate measure, but the response rate was too low to include in the analysis (Engzell & Jonsson, 2015). Low response rates from parents on this topic are not uncommon, with Engzell and Jonsson (2015) finding that in parent samples 'missing data are almost entirely due to unit nonresponse: failure to respond altogether' (p. 317). The findings within our study aligns with this pattern. Therefore, future studies should strive for better gender balance in their recruitment of IB and non-IB samples and gather a range of socio-cultural and economic background indicators to ensure that the propensity score matching procedure includes all relevant variables.

It is important to acknowledge that due to the COVID-19 outbreak, there were difficulties in accessing IB schools during the initially planned targeted authorisation year, which was set to be no later than 2017. Additionally, although we were generally successful in recruitment, it was challenging to recruit Grade 9 students from non-IB schools in Norway, which was originally part of our research design. One school had to withdraw due to pandemic-related constraints, which also contributed to the lack of representation from this year's level. From a methodological standpoint, the matching process was less successful for Norway than for Australia and England, which may be due to the smaller number of Grade 9 students in the non-IB sample in Norway. To improve the matching procedure, it is necessary to ensure that all sub-groups are adequately represented across different grades and countries. In addition, some Norwegian schools in our sample had only recently implemented the MYP (e.g., 2019), possibly with lower levels of fidelity, due to the disruptions caused by COVID-19. Therefore, our findings may be explained by either a less effective matching process for the Norwegian cohort or the fidelity of the MYP implementation in Norway.

In the present study, the two groups were pre-existing (MYP and non-MYP), determined by participants' choices to engage in different programmes. Consequently, the presence of self-selection and the absence of non-random allocation among participants introduce potential biases between these groups. Despite this limitation, the outcomes remain valuable in generating hypotheses concerning probable disparities in critical thinking between MYP and non-MYP groups in this study. Moreover, within each group, our sampling approach ensured that students had an equal opportunity to participate. Additionally, the MYP demonstrates clear strengths, particularly evident in the moderate to large effect sizes observed, especially in the context of England. However, future researchers should consider implementing a design that incorporates true randomisation in assigning students to each curricula type (Stuart, 2010). By closely observing the critical thinking development of the

participants under these controlled conditions, the conclusions that one can draw from the findings can be enhanced.

Finally, critical thinking, as a multidimensional construct, has been previously explored in terms of critical thinking disposition and practice, which has been argued to be different to critical thinking skills (Giancarlo et al., 2004; Tunjungsari & Takwin, 2021). Though the dispositional element of critical thinking was not assessed in our current study, there are instances in each of the three curricula studied where dispositional elements of critical thinking appear, albeit not explicitly. Moreover, there is a lack of studies on critical dispositions across each context, with most curricula focusing on the cognitive (skill) dimension of critical thinking (Paul & Elder, 2012). For example, peer-reviewed studies on critical thinking (skill and disposition) in England's secondary-school classrooms are limited with only 16 peer-reviewed studies published between 2012 and 2021 (see appendix 2 in Bhatt, 2022), and none focused explicitly on critical thinking disposition. Therefore, in addition to expanding research in this area, it is recommended that an approach to examining differences in the skill versus dispositional elements of critical thinking between IB and non-IB samples should be explored to offer a more nuanced understanding of each curriculum. Although explicit teaching is crucial for the development of critical thinking skills, there is still a requirement for strong methodological designs that can compare different approaches to teaching critical thinking disposition, such as explicit, immersion, infusion and mixed methods. These methods should be tested in various configurations and intensities and at different grade levels of secondary education across various educational curricula.

Concluding remarks

Given the broader findings of the MYP's advantage, this study offers timely evidence of a useful curricula for advancing critical thinking among secondary-aged students, particularly as PISA 2018 found that only 9% of 15-year-olds can effectively distinguish fact from opinion (OECD, 2021). Therefore, we recommend the expansion of MYP instructional approaches within the classroom, including explicit instruction, self-directed learning, explicit and learning transfer. The recommendation for explicit instruction is also valued within the Norwegian context, and there is more work to be done to achieve a more unified work on the implementation of critical thinking in England's and Australia's national curricula. Not emphasising critical thinking as part of the educational goal overall could mean that its understanding and implementation among teachers is less effective than it could otherwise be. Alongside this, future research is needed to evaluate which MYP instructional approaches are more effective for students' different critical thinking components and outcomes (e.g., induction, assumption) and in what context. This will help strengthen the specificity of classroom interventions and pre-service teacher training aimed at bolstering students' critical thinking skills.

FUNDING INFORMATION

This project was funded by a grant from the International Baccalaureate (IB) Organisation. The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript. The funders provided feedback on the final report submitted as part of the larger project within which this study is situated. However, the funders had no direct role in the analysis or writing of this paper.

CONFLICT OF INTEREST STATEMENT

No potential conflict of interest was reported by the author(s).

DATA AVAILABILITY STATEMENT

Research data are not shared.

ETHICS STATEMENT

This study has been reviewed by, and received ethics clearance through, the University of Oxford's Central University Research Ethics Committee, [Ethics Approval Number: CIA-21-286].

ENDNOTES

[1] In Australia, lower secondary school is for students from 12 to 16 years old, covering Years 7 to 10. In England, lower secondary school is for students from 12 to 14 years old, covering Years 7 to 9. In Norway, lower secondary school (*ungdomsskole*) is for students from 13 to 16 years old, covering Years 8 to 10.

[2] In the current study, non-MYP refers to the national curriculum. It does not refer to the other programmes (e.g., PYP, DP, Career Programme [CP]) offered by the IB.

[3] The mean and standard deviation refer to the duration of time that MYP students have been enrolled in the programme.

REFERENCES

Ab Kadir, M. A. (2018). An inquiry into critical thinking in the Australian curriculum: Examining its conceptual understandings and their implications on developing critical thinking as a "general capability" on teachers' practice and knowledge. *Asia Pacific Journal of Education*, *38*(4), 533–549. https://doi.org/10.1080/02188791.2018.1535424

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research*, *85*(2), 275–314. https://doi.org/10.3102/0034654314551063

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

Acevedo, E. C., & Hess, C. (2022). The link between critical thinking and personality: Individual differences in a concern for truth. *Modern Psychological Studies*, *27*(1), 9. https://scholar.utc.edu/mps/vol27/iss1/9

Akpur, U. (2020). Critical, reflective, creative thinking and their reflections on academic achievement. *Thinking Skills and Creativity*, *37*, 1–8. https://doi.org/10.1016/j.tsc.2020.100683

Aktoprak, A., & Hursen, C. (2022). A bibliometric and content analysis of critical thinking in primary education. *Thinking Skills and Creativity*, *44*, 101029. https://doi.org/10.1016/J.TSC.2022.101029

Alazzi, K. F. (2008). Teachers' perceptions of critical thinking: A study of Jordanian secondary school social studies teachers. *The Social Studies*, *99*(6), 243–248. https://doi.org/10.3200/TSSS.99.6.243-248

Alwehaibi, H. (2012). Novel programme to promote critical thinking among higher education students: Empirical study from Saudi Arabia. *Asian Social Science*, *8*(11), 193–204. https://doi.org/10.5539/ass.v8n11p193

Anderson, A. D., Hunt, A. N., Powell, R. E., & Dollar, C. B. (2013). Student perceptions of teaching transparency. *Journal of Effective Teaching*, *13*(2), 38–47.

Aoki, S. (2018). *The effects of critical thinking in international baccalaureate classes in Japan*. (Doctoral dissertation). Union University.

Archila, P. A., Molina, J., Danies, G., Truscott De Mejía, A.-M., & Restrepo, S. (2022). Using the controversy over human race to introduce students to the identification and the evaluation of arguments. *Science & Education*, *31*, 861–892. https://doi.org/10.1007/s11191-021-00299-8

Arı, D. (2020). *The effect of skills-based critical thinking education on critical thinking skills of 4th grade students*. [Unpublished Master's thesis]. Çukurova University.

Arısoy, B., & Aybek, B. (2021). The effects of subject-based critical thinking education in mathematics on students' critical thinking skills and virtues. *Eurasian Journal of Educational Research*, *92*, 99–120. https://doi.org/10.14689/ejer.2021.92.6

Ateşkan, A., Dulun, Ö., & Lane, J. F. (2016). *Middle years Programme (MYP) implementation in Turkey*. International Baccalaureate Organization.

Australian Council for Educational Research [ACER]. (2015). *The international baccalaureate (IB) middle years Programme (MYP): Comparing IB diploma Programme outcomes of students who complete the MYP and other middle years courses of study*. International Baccalaureate Organization.

Australian Curriculum, Assessment and Reporting Authority. (2014). *Creative and critical thinking learning continuum*. https://www.australiancurriculum.edu.au/media/1072/general-capabilities-creative-and-critical-thinking-learning-continuum.pdf

Australian Curriculum, Assessment and Reporting Authority. (n.d.). *Critical and creative thinking*. https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/critical-and-creative-thinking/

Beckwitt, A., Van Camp, D., & Carter, J. (2015). *International baccalaureate implementation study: Examination of district-wide implementation in the US* (Research Brief). International Baccalaureate Organization.

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), 25–42. https://doi.org/10.19030/tlc.v8i2.3554

Belas, O. (2019). Knowledge, the curriculum, and democratic education: The curious case of school English. *Research in Education*, *103*(1), 49–67. https://doi.org/10.1177/0034523719839095

Bezanilla, M. J., Fernández-Nogueira, D., Poblete, M., & Galindo-Domínguez, H. (2019). Methodologies for teaching-learning critical thinking in higher education: The teacher's view. *Thinking Skills and Creativity*, *33*, 1–10. https://doi.org/10.1016/j.tsc.2019.100584

Bhatt, P. M. (2022). *What are teacher perceptions of critical thinking and how do they develop the critical thinking of the pupils they teach? A case study from an inner London secondary school* (Doctoral dissertation). University of Leicester.

Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2014). Measuring fidelity and adaptation: Reliability of a instrument for school-based prevention programs. *Evaluation & the Health Professions*, *37*(2), 231–257. https://doi.org/10.1177/0163278713476882

British Council. (2016). *Unlocking a world of potential: Core skills for learning, work and society*. https://schoolsonline.britishcouncil.org/sites/default/files/core_skills_brochure_britishcouncil_2016-17_1.pdf

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

Calnin, G., Waterson, M., Richards, S., & Fisher, D. (2018). Developing leaders for international baccalaureate world schools. *Journal of Research in International Education*, *17*(2), 99–115. https://doi.org/10.1177/1475240918790703

Chaffee, J. (1994). *Thinking critically*. Houghton Mifflin.

Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018). Extending cognitive load theory to incorporate working memory resource depletion: Evidence from the spacing effect. *Educational Psychology Review*, *30*, 483–501. https://doi.org/10.1007/s10648-017-9426-2

Chirgwin, S. K., & Huijser, H. (2015). Cultural variance, critical thinking, and indigenous knowledges: Exploring a both-ways approach. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 335–350). Palgrave Macmillan. https://doi.org/10.1057/9781137378057.0027

Clifford, J. S., Boufal, M. M., & Kurtz, J. E. (2004). Personality traits and critical thinking skills in college students: Empirical tests of a two-factor theory. *Assessment*, *11*(2), 169–176. https://doi.org/10.1177/1073191104263250

Cole, D. R., Ullman, J., Gannon, S., & Rooney, P. (2015). Critical thinking skills in the international Baccalaureate's "theory of knowledge" subject: Findings from an Australian study. *Australian Journal of Education*, *59*(3), 247–264. https://doi.org/10.1177/0004944115603529

D'Alessio, F. A., Avolio, B. E., & Charles, V. (2019). Studying the impact of critical thinking on the academic performance of executive MBA students. *Thinking Skills and Creativity*, *31*, 275–283. https://doi.org/10.1016/j.tsc.2019.02.002

Davies, M., & Barnett, R. (2015). The palgrave handbook of critical thinking in higher education. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education*. Palgrave Macmillan. https://doi.org/10.1057/9781137378057

Davies, W. M. (2006). An 'infusion' approach to critical thinking: Moore on the critical thinking debate. *Higher Education Research & Development*, *25*(2), 179–193. https://doi.org/10.1080/07294360600610420

Department for Education. (2013a). *National curriculum in England: Framework document*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/381344/Master_final_national_curriculum_28_Nov.pdf

Department for Education. (2013b). *Teachers' standards guidance for school leaders, school staff and governing bodies*. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/665520/Teachers__Standards.pdf

Department for Education. (2015). *Policy Paper – 2010 to 2015 government policy: School and college qualifications and curriculum*. https://www.gov.uk/government/publications/2010-to-2015-government-policy-school-and-college-qualifications-and-curriculum/2010-to-2015-government-policy-school-and-college-qualifications-and-curriculum

Dever, R. (2019). The international baccalaureate middle years programme (MYP) in US middle level schools. In D. C. Virtue (Ed.), *International handbook of middle level education theory, research, and policy* (pp. 31–40). Routledge.

Dewey, J. (1909). *How we think*. D.C. Heath and Co.

Dickson, A., Perry, L., & Ledger, S. (2021). Challenges of the international baccalaureate middle years Programme: Insights for school leaders and policy makers. *Education Policy Analysis Archives*, *29*(137), 1–28. https://doi.org/10.14507/epaa.29.5630

Dickson, A., Perry, L. B., & Ledger, S. (2018). Impacts of international baccalaureate programmes on teaching and learning: A review of the literature. *Journal of Research in International Education*, *17*(3), 240–261. https://doi.org/10.1177/1475240918815801

Dickson, A., Perry, L. B., & Ledger, S. (2020). Challenges impacting student learning in the international baccalaureate middle years programme. *Journal of Research in International Education*, *19*(3), 183–201. https://doi.org/10.1177/1475240920976228

Diepeveen, S., & Pinet, M. (2022). User perspectives on digital literacy as a response to misinformation. *Development Policy Review*, *40*, e12671. https://doi.org/10.1111/dpr.12671

Dilekçi, A., & Karatay, H. (2023). The effects of the 21st century skills curriculum on the development of students' creative thinking skills. *Thinking Skills and Creativity*, *47*, 1–11. https://doi.org/10.1016/j.tsc.2022.101229

Donnelly, K., & Wiltshire, K. (2014). *Review of the Australian curriculum: Final report*. Commonwealth of Australia.

Double, K. S., El Masri, Y., McGrane, J. A., & Hopfenbeck, T. N. (2023). Do IB students have higher critical thinking? A comparison of IB with national education programs. *Thinking Skills and Creativity*, *101416*, 1–9. https://doi.org/10.1016/j.tsc.2023.101416

Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education*, *46*(3), 333–345. https://doi.org/10.1080/0305764X.2016.1142504

Dworak, E. M., Revelle, W., Doebler, P., & Condon, D. M. (2021). Using the international cognitive ability resource as an open-source tool to explore individual differences in cognitive ability. *Personality and Individual Differences*, *169*, 109906. https://doi.org/10.1016/j.paid.2020.109906

Dwyer, C. P. (2017). *Critical thinking: Conceptual perspectives and practical guidelines*. Cambridge University Press.

Engzell, P. (2021). What do books in the home proxy for? A cautionary tale. *Sociological Methods & Research*, *50*(4), 1487–1514. https://doi.org/10.1177/0049124119826143

Engzell, P., & Jonsson, J. O. (2015). Estimating social and ethnic inequality in school surveys: Biases from child misreporting and parent nonresponse. *European Sociological Review*, *31*(3), 312–325. https://doi.org/10.1093/esr/jcv005

Ennis, R. (1991). Critical thinking. *Teaching Philosophy*, *14*(1), 4–18. https://doi.org/10.5840/inquiryctnews201126215

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, *18*(3), 4–10. https://doi.org/10.3102/0013189X018003004

EnniEnniss, R. H. (1990). The extent to which critical thinking is subject-specific: Further clarification. *Educational Researcher*, *19*(4), 13–16. https://doi.org/10.3102/0013189X019004013

Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice*, *32*(3), 179–186. https://doi.org/10.1080/00405849309543594

Ennis, R. H. (2013). Critical thinking across the curriculum: The wisdom CTAC program. *Inquiry: Critical Thinking Across the Disciplines*, *28*(2), 25–45. https://doi.org/10.5840/inquiryct20132828

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 test* (Fifth ed.). The Critical Thinking Company.

Ennis, R. H., & Weir, E. E. (1985). *The Ennis-Weir critical thinking essay test: An instrument for teaching and testing*. Midwest Publications.

Erikson, M. G., & Erikson, M. (2019). Learning outcomes and critical thinking–good intentions in conflict. *Studies in Higher Education*, *44*(12), 2293–2303. https://doi.org/10.1080/03075079.2018.1486813

Eriksson, K., Lindvall, J., Helenius, O., & Ryve, A. (2022). Higher-achieving children are better at estimating the number of books at home: Evidence and implications. *Frontiers in Psychology*, *13*, 1026387. https://doi.org/10.3389/fpsyg.2022.1026387

Evangelisto, C. (2021). Critical thinking in STEM: A qualitative study of community college teaching techniques. *Journal of STEM Education*, *22*(2), 33–39.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (Executive summary: "The Delphi report")*. California Academic Press. http://assessment.aas.duke.edu/documents/Delphi_Report.pdf

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

French, B. F., Hand, B., Therrien, W. J., & Valdivia Vazquez, J. A. (2012). Detection of sex differential item functioning in the Cornell critical thinking test. *European Journal of Psychological Assessment*, *28*(3), 201–207. https://doi.org/10.1027/1015-5759/a000127

Fuad, N. M., Zubaidah, S., Mahanal, S., & Suarsini, E. (2017). Improving junior high schools' critical thinking skills based on test three different models of learning. *International Journal of Instruction*, *10*(1), 101–116. https://doi.org/10.12973/iji.2017.1017a

Gann, D. A. (2013). A few considerations on critical thinking instruction. Considerations-On-Criticalthinking-İnstruction-David-A-Gann/Erişim Tarihi, 24, 2014.

Giancarlo, C. A., Blohm, S. W., & Urdan, T. (2004). Assessing secondary students' disposition toward critical thinking: Development of the California measure of mental motivation. *Educational and Psychological Measurement*, *64*(2), 347–364. https://doi.org/10.1177/0013164403258464

Gilbert, R. (2019). General capabilities in the Australian curriculum: Promise, problems and prospects. *Curriculum Perspectives*, *39*, 169–177. https://doi.org/10.1007/s41297-019-00079-z

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34. https://doi.org/10.1037/0003-066X.48.1.26

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Gonski, D., Arcus, T., Boston, K., Gould, V., Johnson, W., O'Brien, L., Perry, L., Perry, L., & Roberts, M. (2018). *Through growth to achievement: Report of the review to achieve educational excellence in Australian schools*. Department of Education and Training.

Goodsett, M. (2020). Best practices for teaching and assessing critical thinking in information literacy online learning objects. *The Journal of Academic Librarianship*, *46*(5), 102163. https://doi.org/10.1016/j.acalib.2020.102163

Grewal, A. (2001). *The MYP and the IGCSE: Compatible or conflicting programmes?: Studying the issues that arise when trying to merge the two programmes into one* (Doctoral dissertation). Oxford Brookes University.

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology Press.

Halpern, D. F., & Butler, H. A. (2018). Is critical thinking a better model of intelligence? In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 183–196). Cambridge University Press. https://doi.org/10.1017/9781316817049.013

Halpern, D. F., & Dunn, D. S. (2021). Critical thinking: A model of intelligence for solving real-world problems. *Journal of Intelligence*, *9*(2), 22. https://doi.org/10.3390/jintelligence9020022

Hamza-Lup, F. G., & Sopin, I. (2008). Haptics and extensible 3D in web-based environments for e- learning and simulation. In J. Cordeiro, J. Filipe, & S. Hammoudi (Eds.), *Proceedings of the 4th international conference on web information systems and technologies* (pp. 309–315). INSTICC Press.

Hansen, B. B. (2004). Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association*, *99*(467), 609–618. https://doi.org/10.1198/016214504000000647

Hansen, W. B., Pankratz, M. M., Dusenbury, L., Giles, S. M., Bishop, D. C., Albritton, J., Albritton, L. P., & Strack, J. (2013). Styles of adaptation: The impact of frequency and valence of adaptation on preventing substance use. *Health Education*, *113*(4), 345–363. https://doi.org/10.1108/09654281311329268

Hargittai, E., Piper, A. M., & Morris, M. R. (2019). From internet access to internet skills: Digital inequality among older adults. *Universal Access in the Information Society*, *18*, 881–890. https://doi.org/10.1007/s10209-018-0617-5

Haynes, A., Lisic, E., Harris, K., Leming, K., Shanks, K., & Stein, B. (2015). Using the critical thinking assessment test (CAT) as a model for designing within-course assessments. *Inquiry: Critical Thinking across the Disciplines*, *30*(3), 38–48. https://doi.org/10.5840/inquiryct201530316

Heppt, B., Olczyk, M., & Volodina, A. (2022). Number of books at home as an indicator of socioeconomic status: Examining its extensions and their incremental validity for academic achievement. *Social Psychology of Education*, *25*, 903–928. https://doi.org/10.1007/s11218-022-09704-8

Hill, I. (2001). Early stirrings: The beginnings of the international education movement. *International Schools Journal*, *20*(2), 11–21.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199–236. https://doi.org/10.1093/pan/mpl013

Holland, E. C. (2016). *" IB Pedia": Informing educators of the international baccalaureate middle years program through collaborative website development*. California State University.

IBO. (2012). *Teaching the disciplines in the MYP: Nurturing big ideas and deep understanding*. International Baccalaureate Organization.

IBO. (2014a). *MYP: From principles into practice*. International Baccalaureate Organization.

IBO. (2014b). *Language and literature guide*. International Baccalaureate Organization.

IBO. (2018). *What is an IB education?* International Baccalaureate Organization.

IBO. (2022). *What is the MYP? Curriculum*. International Baccalaureate Organization.

IBO. (2023). *Key facts about the MYP*. International Baccalaureate Organization.

Indrašienė, V., Jegelevičienė, V., Merfeldaitė, O., Penkauskienė, D., Pivorienė, J., Railienė, A., Sadauskas, J., & Valavičienė, N. (2021). *Critical thinking in higher education and labour market*. Peter Lang International Academic Publishers. https://doi.org/10.3726/18636

Joshi, S. C., Gupta, K., & Manektala, S. (2022). Misinformation, public opinion, and the role of critical thinking. *International Journal of Management and Humanities*, *8*(9), 15–18. https://doi.org/10.35940/ijmh.i1483.058922

Kaklanis, N., Votis, K., & Tzovaras, D. (2015). Adding haptic feedback to web applications towards improving end-users' cognitive capabilities. Paper presented at the 6th IEEE international conference on cognitive Infocommunications (CogInfoCom), Hungary.

Ku, K. Y., Ho, I. T., Hau, K. T., & Lai, E. C. (2014). Integrating direct and inquiry-based instruction in the teaching of critical thinking: An intervention study. *Instructional Science*, *42*, 251–269. https://doi.org/10.1007/s11251-013-9279-0

Lai, E. R. (2011). *Critical thinking: A literature review*. Pearson.

Livingstone, S., Ólafsson, K., Helsper, E. J., Lupiáñez-Villanueva, F., Veltri, G. A., & Folkvord, F. (2017). Maximizing opportunities and minimizing risks for children online: The role of digital skills in emerging strategies of parental mediation. *Journal of Communication*, *67*(1), 82–105. https://doi.org/10.1111/jcom.12277

Lombardi, L., Mednick, F. J., De Backer, F., & Lombaerts, K. (2021). Fostering critical thinking across the primary school's curriculum in the European schools system. *Education Sciences*, *11*(9), 1–19. https://doi.org/10.3390/educsci11090505

Manalo, E., Kusumi, T., Koyasu, M., Michita, Y., & Tanaka, Y. (2015). Do students from different cultures think differently about critical and other thinking skills? In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 299–316). Palgrave Macmillan. https://doi.org/10.1057/9781137378057_19

Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*, *6*(1), 1–13. https://doi.org/10.1016/j.tsc.2010.08.002

McGrane, J. A., Johnston, S.-K., Morancho, M. V., & Hopfenbeck, T. N. (2022). *Critical thinking development in the IB middle years programme: An international multisite evaluation*. International Baccalaureate Organization.

Meirbekov, A., Maslova, I., & Gallyamova, Z. (2022). Digital education tools for critical thinking development. *Thinking Skills and Creativity*, *44*, 1–8. https://doi.org/10.1016/j.tsc.2022.101023

Michael, J. J., Devaney, R. L., & Michael, W. B. (1980). The factorial validity of the Cornell critical thinking test for a junior high school sample. *Educational and Psychological Measurement*, *40*(2), 437–450. https://doi.org/10.1177/001316448004000223

Mullis, I. V. S., Martin, M. O., & Liu, J. (2019). *PIRLS 2021 assessment framework*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, IEA. https://pirls2021.org/frameworks/wp-content/uploads/sites/2/2019/04/P21_FW_Ch2_Questionnaires.pdf

Munkebye, E., & Gericke, N. (2022). Primary school teachers' understanding of critical thinking in the context of education for sustainable development. In B. Puig & M. P. Jimenez-Aleixandre (Eds.), *Critical thinking in biology and environmental education: Facing challenges in a post-truth world* (pp. 249–266). Springer.

Musi, E., Carmi, E., Reed, C., Yates, S., & O'Halloran, K. (2023). Developing misinformation immunity: How to reason-check fallacious news in a human–computer interaction environment. *Social Media + Society*, *9*(1), 1–18. https://doi.org/10.1177/20563051221150407

Norwegian Directorate for Education and Training. (2020). *Curriculum for Natural science* (*NAT01-04*). https://www.udir.no/lk20/nat01-04?lang=eng

Norwegian Ministry of Education and Research. (2017). *Core curriculum for primary, lower secondary and upper secondary education: Introduction*. https://www.regjeringen.no/contentassets/53d21ea2bc3a4202b86b83cfe82da93e/core-curriculum.pdf

NSW Board of Studies, Teaching and Educational Standards. (2014). *Submission to the Australian Government review of the Australian curriculum*. http://www.boardofstudies.nsw.edu.au/

OECD. (2021). *21st-century readers: Developing literacy skills in a digital world*. OECD Publishing.

Ofqual. (2017). *GCE qualification level conditions and requirements*. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/644330/gce-qualification-level-conditions-and-requirements.pdf

Orhan, A., & Çeviker Ay, Ş. (2022). How to teach critical thinking: An experimental study with three different approaches. *Learning Environments Research*, *26*, 199–217. https://doi.org/10.1007/s10984-022-09413-1

Ott, A. (2019). Kritisk tenkning og bærekraft i fagfornyelsen [Critical thinking and sustainability in the curriculum]. In M. Ferrer & A. Wetlesen (Eds.), *Kritisk tenkning i samfunnsfag [Critical thinking in social science]* (pp. 30–44). Universitetsforlaget.

Oxford, Cambridge and RSA Examinations (OCR). (2011). *Level 2 award thinking and reasoning skills specification version 2*. http://www.ocr.org.uk/qualifications/other-general-qualifications-thinking-and-reasoningskills-level-2-award-j930/

Oxford, Cambridge and RSA Examinations (OCR). (2016). *Level 2 award in thinking and reasoning skills (J930) (QAN 500/5687/6) – withdrawal of level 2 award in thinking and reasoning skills (J930) (QAN 500/5687/6)*. https://www.ocr.org.uk/administration/support-and-tools/subject-updates/think-and-reasoning-skills-l2-award-150416/

Paul, R., & Elder, L. (2012). *Critical thinking: Tools for taking charge of your learning and your life* (3rd ed.). Pearson Education.
Paul, R., & Elder, L. (2019). *A guide for educators to critical thinking competency standards: Standards, principles, performance indicators, and outcomes with a critical thinking master rubric*. The Foundation for Critical Thinking. www.criticalthinking.org
Pendergast, D., Dole, S., & Rentoule, D. (2014). Teacher dependent discourse patterns: Variations in student discourse roles in an international baccalaureate school-a mathematics focus. *The International Schools Journal*, *33*(2), 16–25. http://hdl.handle.net/10072/63572
Perry, L., Ledger, S., & Dickson, A. (2018). *What are the benefits of the International Baccalaureate Middle Years Programme for teaching and learning? Perspectives from stakeholders in Australia*. International Baccalaureate Organization.
Pnevmatikos, D., Christodoulou, P., & Georgiadou, T. (2019). Promoting critical thinking in higher education through the values and knowledge education (VaKE) method. *Studies in Higher Education*, *44*(5), 892–901. https://doi.org/10.1080/03075079.2019.1586340
Queensland Curriculum and Assessment Authority (QCAA). (2015). *21st century skills for senior education: An analysis of educational trends*. QCAA.
R Core Team. (2013). *A language and environment for statistical computing R Foundation for Statistical Computing, Austria*. http://www.R-project.org/
Robertson, J. E. (2011). *An analysis of the role of international baccalaureate middle years Programme coordinator* (Doctoral dissertation). University of Bath.
Rodríguez-de-Dios, I., van Oosten, J. M., & Igartua, J. J. (2018). A study of the relationship between parental mediation and adolescents' digital skills, online risks and online opportunities. *Computers in Human Behavior*, *82*, 186–198. https://doi.org/10.1016/j.chb.2018.01.012
Santos, L. F., Gyenes, A., Fujisaki, S., & Gay, S. (2021). Critical thinking viii(1).
Scheie, E., Haug, B. S., & Erduran, S. (2022). Critical thinking in the Norwegian science curriculum. *Acta Didactica Norden (ADNO)*, *16*(2), 1–27. https://doi.org/10.5617/adno.9060
Siregar, N. R. (2021). Explicit instruction and executive functioning capacity: A new direction in cognitive load theory. *Journal of Education*, *203*(2), 1–8. https://doi.org/10.1177/00220574211033256
Sizmur, J., & Cunningham, R. (2013). *International baccalaureate middle years Programme (MYP) in the UK*. National Foundation for Educational Research.
Skourdoumbis, A. (2016). Articulations of teaching practice: A case study of teachers and 'general capabilities'. *Asia Pacific Education Review*, *17*(4), 545–554. https://doi.org/10.1007/s12564-016-9460-7
Smith, T. E., Rama, P. S., & Helms, J. R. (2018). Teaching critical thinking in a GE class: A flipped model. *Thinking Skills and Creativity*, *28*, 73–83. https://doi.org/10.1016/j.tsc.2018.02.010
Soneral, P. A. G., & Wyse, S. A. (2017). A SCALE-UP mock-UP: Comparison of student learning gains in high- and low-tech active-learning environments. *CBE Life Sciences Education*, *16*(1), 1–15. https://doi.org/10.1187/cbe.16-07-0228
Sperandio, J. (2010). School program selection: Why schools worldwide choose the international baccalaureate middle years program. *Journal of School Choice*, *4*(2), 137–148. https://doi.org/10.1080/15582159.2010.483916
Stanovich, K. E., & West, R. F. (2008). On the failure of cognitive ability to predict myside and one-sided thinking biases. *Thinking & Reasoning*, *14*(2), 129–167. https://doi.org/10.1080/13546780701679764
Stevenson, H., Shah, S., Bailey, L., Cooker, L., Winchip, E., & Karak, M. (2017). *The international baccalaureate middle years Programme (MYP) implementation in The United Arab Emirates*. International Baccalaureate Organisation.
Storz, M., & Hoffman, A. (2018). Becoming an international baccalaureate middle years program: Perspectives of teachers, students, and administrators. *Journal of Advanced Academics*, *29*(3), 216–248. https://doi.org/10.1177/1932202X18770171
Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *25*(1), 1–21. https://doi.org/10.1214/09-STS313
Stuart, E. A., & Green, K. M. (2008). Using full matching to estimate causal effects in nonexperimental studies: Examining the relationship between adolescent marijuana use and adult outcomes. *Developmental Psychology*, *44*(2), 395–406. https://doi.org/10.1037/0012-1649.44.2.395
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5
Taylor, M. L., & Porath, M. (2006). Reflections on the international baccalaureate program: Graduates' perspectives. *Journal of Secondary Gifted Education*, *17*(3), 149–158. https://doi.org/10.4219/jsge-2006-398
Termaat, A. (2023). Nurturing interdisciplinary practice in small secondary schools. *Cogent Education*, *10*(2), 2225003. https://doi.org/10.1080/2331186X.2023.2225003
Thomas, K., & Lok, B. (2015). Teaching critical thinking: An operational framework. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education*. Palgrave Macmillan.

Tiruneh, D. T., De Cock, M., Weldeslassie, A. G., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics: Development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education*, *15*(4), 663–682. https://doi.org/10.1007/s10763-016-9723-0

Tiruneh, D. T., Weldeslassie, A. G., Kassa, A., Tefera, Z., De Cock, M., & Elen, J. (2016). Systematic design of a learning environment for domain-specific and domain-general critical thinking skills. *Educational Technology Research and Development*, *64*(3), 481–505. https://doi.org/10.1007/s11423-015-9417-2

Tunjungsari, H., & Takwin, B. (2021). Understanding critical thinking practice in everyday life through stages of critical thinking skills and disposition. *Mind, Brain, and Education*, *15*(3), 225–231. https://doi.org/10.1111/mbe.12279

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. *SAGE Open*, *10*(1), 1–14. https://doi.org/10.1177/2158244019900176

Vardi, I. (2015). The relationship between self-regulation, personal epistemology, and becoming a "critical thinker": Implications for pedagogy. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 197–212). Pagrave Macmillan. https://doi.org/10.1057/9781137378057_13

Velez, G., & Power, S. A. (2020). Teaching students how to think, not what to think: Pedagogy and political psychology. *Journal of Social and Political Psychology*, *8*(1), 388–403. https://doi.org/10.5964/jspp.v8i1.1284

Ventista, O. M. (2018). Multi-trait multi-method matrices for the validation of creativity and critical thinking assessments for secondary school students in England and Greece. *International Journal of Assessment Tools in Education*, *5*(1), 15–32. https://doi.org/10.21449/ijate.335167

Victorian Curriculum and Assessment Authority (VCAA). (2015). *Victorian curriculum F–10: Revised curriculum planning and reporting guidelines*. VCAA.

Vieira, R. M., Tenreiro-Vieira, C., & Martins, I. P. (2011). Critical thinking: Conceptual clarification and its importance in science education. *Science Education International*, *22*(1), 43–54. https://eric.ed.gov/?id=EJ941655

Visser, A. (2010). International education in a national context: Introducing the international baccalaureate middle years Programme in Dutch public schools. *Journal of Research in International Education*, *9*(2), 141–152. https://doi.org/10.1177/1475240910370816

Volman, M., & ten Dam, G. (2015). Critical thinking for educated citizenship. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 593–603). Palgrave Macmillan. https://doi.org/10.1057/9781137378057_35

Wade, J., Wolanin, N., & McGaughey, T. (2016). *A study of critical thinking skills in the international baccalaureate middle years Programme*. The International Baccalaureate Organization.

Wahyudi, Verawati, N. N. S. P., Ayub, S., & Prayogi, S. (2019). Effectiveness of inquiry-creative-process learning model to promote critical thinking ability of prospective physics teachers. *Journal of Physics: Conference Series*, *1417*(1), 1–6. https://doi.org/10.1088/1742-6596/1417/1/012071

Walker, A., & Kettler, T. (2020). Developing critical thinking skills in high ability adolescents: Effects of a debate and argument analysis curriculum. *Talent*, *10*(1), 21–39. https://doi.org/10.46893/talent.758473

Wallace, E. D., & Jefferson, R. N. (2015). Developing critical thinking skills: Assessing the effectiveness of workbook exercises. *Journal of College Teaching & Learning*, *12*(2), 101–108. https://doi.org/10.19030/tlc.v12i2.9187

Waterbury, S. M. (2018). *Icing the batter: The MYP coordinator, principal leadership, and school reform* (Unpublished doctoral dissertation). Georgia State University.

Watt, M. G. (2006). *Looking at curriculum change in Tasmania: Will essential learnings promote successful reform?* Online Submission.

Willingham, D. T. (2008). Critical thinking: Why is it so hard to teach? *Arts Education Policy Review*, *109*(4), 21–32. https://doi.org/10.3200/AEPR.109.4.21-32

Wolanin, N., & Wade, J. (2012). *International baccalaureate middle years program: Teacher perceptions*. International Baccalaureate.

Wootten, L. M. (2019). *Exploring 21st century teaching and learning skills in the international baccalaureate continuum training and practice* (Doctoral dissertation). Northcentral University.

Wright, E., Lee, M., Tang, H., & Chak Pong Tsui, G. (2016). Why offer the international baccalaureate middle years Programme? A comparison between schools in Asia-Pacific and other regions. *Journal of Research in International Education*, *15*(1), 3–17. https://doi.org/10.1177/1475240916635896

Xu, E., Wang, W., & Wang, Q. (2023). The effectiveness of collaborative problem solving in promoting students' critical thinking: A meta-analysis based on empirical literature. *Humanities and Social Sciences Communications*, *10*(1), 1–11. https://doi.org/10.1057/s41599-023-01508-1

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*, 1–15. https://doi.org/10.1016/j.tsc.2022.101002

Zandvakili, E., Washington, E., Gordon, E. W., Wells, C., & Mangaliso, M. (2019). Teaching patterns of critical thinking: The 3CA model—Concept maps, critical thinking, collaboration, and assessment. *SAGE Open*, *9*(4), 1–15. https://doi.org/10.1177/2158244019885142

Zhang, Q., Tang, H., & Xu, X. (2022). Analyzing collegiate critical thinking course effectiveness: Evidence from a quasi-experimental study in China. *Thinking Skills and Creativity*, *45*, 1–14. https://doi.org/10.1016/j.tsc.2022.101105

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Johnston, S.-K., McGrane, J. A., Vendrell-Morancho, M., & Hopfenbeck, T. N. (2023). A multi-country comparison of lower secondary students' critical thinking under the International Baccalaureate and national curricula. *Review of Education*, *11*, e3442. https://doi.org/10.1002/rev3.3442