City Research Online

City St George's, University of London

Citation: Tan, A. J. Y., Davies, J. L., Nicolson, R. I. & Karaminis, T. (2023). Learning critical thinking skills online: can precision teaching help?. Educational technology research and development, 71(3), pp. 1275-1296. doi: 10.1007/s11423-023-10227-y

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/32230/

Link to published version: https://doi.org/10.1007/s11423-023-10227-y

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

* This version of the article has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at:
https://link.springer.com/article/10.1007/s11423-023-10227-y#Sec1
Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use
https://www.springernature.com/gp/open-research/policies/acceptedmanuscript-terms

Learning critical thinking skills online: Can Precision Teaching help?

Angel J. Y. Tan[1], Jean L. Davies[2], Roderick I. Nicolson[2], & Themis Karaminis[2]

[1] Department of Psychology, Birmingham City University, Birmingham, UK

[2] Department of Psychology, Edge Hill University, Ormskirk, UK

Author Note

Angel J. Y. Tan, angel.tan@bcu.ac.uk

Jean L. Davies, daviesj@edgehill.ac.uk

Rod I. Nicolson, rod.nicolson@edgehill.ac.uk

Themis Karaminis, themis.karaminis@edgehill.ac.uk

Correspondence concerning this article should be addressed to Angel Tan, Department of Psychology, Birmingham City University, Birmingham, UK. Email: angel.tan@bcu.ac.uk

Abstract

Critical thinking is identified as a key educational outcome in higher education curricula; however, it is not trivial to support students in building this multifaceted ability. In this study, we evaluated a brief online learning intervention focusing on informal fallacy identification, a hallmark critical-thinking skill. The intervention used a bite-sized video learning approach, which has been shown to promote student engagement. Video-based learning was implemented within a Precision Teaching (PT) framework, which modulates the exposure of individual learners to the learning material to enable them to build 'fluency' in the targeted skills. In one of the learning conditions, PT was applied synergistically with domain-general problem-based training to support generalisation. The intervention consisted of two learning episodes and was administered to three groups (learning conditions) of 19 participants each: a PT fluency-based training group; a PT+ group, where PT was combined with problem-based training; and a self-directed learning control group. All three groups showed comparable improvements in fallacy identification on taught (post-episode tests) and unseen materials (post-intervention assessment), with lower-scoring participants showing higher gains than high-scoring participants. The results on the knowledge retention tests a week later were also comparable between groups. Importantly, in the domain-general fallacy-identification assessment (post-intervention), the two PT groups showed higher improvements than the control group. These findings suggest that the integration of bite-sized video learning technologies with PT can improve students' critical-thinking skills. Furthermore, PT, on its own or combined with problem-based training, can improve their ability to generalise learning to novel contexts. We discuss the educational implications of our findings.

Keywords: critical thinking, precision teaching, problem-based learning, video-based learning, computer-assisted learning

Introduction

Critical thinking can be described as the "purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation, and inference, as well as explanations of the considerations on which that judgement is based" (Abrami et al., 2015, p.275). This high-level skill enables individuals to think logically, make appropriate decisions, and solve problems effectively (Peter, 2012). Critical thinking has been associated with academic achievements, enhanced employability, higher financial status, and better real-life decisions (Butler, Pentoney, & Bong, 2017; Facione & Facione, 2001; Hart Research Associates, 2015). It has also been identified as an important educational goal for Higher Education (HE) preparing students for the demands of the 21st Century workplace (Hatcher, 2011; Joynes, Rossignoli, & Amonoo-Kuofi, 2019) and is often included in learning outcomes and assessment standards across disciplines (Forbes, 2018).

However, despite the emphasis that HE curricula place on critical thinking, students present difficulties in demonstrating critical-thinking skills (Harrington et al., 2006; Kreth, Crawford, Taylor, & Brockman, 2010). From educators' perspective, formal training in critical thinking instruction is rarely provided (Broadbear, 2003; Scriven & Paul, 2007), and there is no clear consensus on how critical thinking should be taught (Abrami et al., 2015). Some researchers have suggested that critical thinking builds on generic metacognitive skills, such as reasoning, interpreting, analysing, and should thus be taught across disciplines (Solon, 2007). Other researchers have argued that critical thinking is context-specific and should be taught within the domains in which it is used (Baker, 2001; Brunt, 2005; McPeck, 1981). Critical thinking education is also challenged by the limited contact time for critical discussion and evaluation of the learning content in conventional teacher-led instructional approaches (Mandernach, 2006; Peter, 2012).

The rapid shift of the HE sector to online teaching during the recent COVID-19 pandemic (WHO, 2020) presented educators with new challenges in teaching critical thinking. For example, online learning relies on students feeling comfortable with using and participating in live discussion boards, online debates and focus groups, and this may pose a barrier to student access and engagement in activities relevant to the application of critical thinking skills (MacKnight, 2000). There is a scarcity of studies on instructional strategies to promote critical thinking in online environments (Guiller et al., 2008; Richardson & Ice, 2010).

In this study, we examined the effectiveness of a technology-enhanced learning intervention for critical thinking administered online to HE students during the second round of COVID-19 restrictions in the UK (early 2021). The intervention combined video-based learning with precision teaching, a behaviourally-grounded teaching approach targeted to build so-called fluency on learnt skills. In addition to this, in one of the learning conditions, precision teaching was combined with context-based training to better support the application of learnt knowledge.

Video-based Learning

Video-based learning enables students to learn outside of the physical classrooms and at their own pace (Syed et al., 2020). It also enables educators to enrich mainstream teaching provision with supplementary material, implement diverse pedagogical strategies (e.g., flipped classroom, blended learning; Yousef, Chatti, & Schroeder, 2014), and meet students' individual learning needs and preferences (Carmichael, Reid, & Karpicke, 2018). There is ample evidence that video-based learning can enhance students' engagement (Stockwell, Stockwell, Cennamo, & Jiang, 2015), academic performance (Salina et al., 2012), and motivation (Hill & Nelson, 2011). There is also evidence that these benefits are maximised when videos of a shorter duration are used (Guo, Kim, & Rubin, 2014).

Bite-sized or micro-videos are designed to chunk information into manageable and digestible pieces, making the learning content more accessible and improving the engagement of students with it (Koh, Gottipati, & Shankararaman, 2018). It has been suggested that bite-sized video learning sessions facilitate active learning (Brame, 2016), as students can rewind and review parts of the videos more easily when videos are available in smaller chunks (Carmichael et al., 2018). High-speed internet and improved functionality of mobile devices have also helped to integrate bite-sized learning into everyday routines and support autonomy in learning (Khong & Kabilan, 2020). However, research on the educational uses of videos has mostly focused on subject-relevant knowledge and practical skills rather than on higher-level skills such as critical thinking (Carmichael et al., 2018).

Precision Teaching (PT)

PT refers to a framework for the systematic self-monitoring of learning (Lindsley, 1997) and the effectiveness of instructional approaches (Kubina & Yurich, 2012). PT can also be used to collect students' learning data and tailor instructional methods to the individual student's performance (Sundhu & Kittles, 2016). PT often obtains evidence of learning by measuring fluency, the combination of accuracy and speed in performing a targeted skill (Kubina & Morrison, 2000, p.89), which is a prerequisite for more advanced skills (Kubina & Morrison, 2000). Within the PT framework, fluency is associated with other learning outcomes, including retention - maintaining good performance after an interval without training, endurance - carrying out a task fluently for long durations, stability - not being affected by distractions, and application - combining basic skills to perform a more complex task (abbreviated as RESA; see Binder, 1996; Kubina & Yurich, 2012).

A commonly used fluency-training approach within the PT framework is frequency building (Kubina & Yurich, 2012). Frequency building uses timed repetition of tasks coupled with performance feedback provided immediately after timed trials (Lokke, Lokke, & Arntzen,

2008). This practice is thought to support the acquisition of the targeted skills in a time-efficient manner (Kubina & Yurich, 2012).

Research has shown that frequency-building techniques can support the acquisition of academic skills, such as reading, handwriting, and numeracy (e.g. Chiesa & Robertson, 2000; Hughes, Beverley, & Whitehead, 2007). There is less extensive evidence on whether and how frequency-building approaches could support the learning of models of complex thinking (Commons, Owens, & Will, 2015), improve fluency in complex concepts, such as logical fallacies (Fox & Ghezzi, 2003) and strengthen domain-general cognitive skills (Cuzzocrea, Murdaca, and Oliva, 2011).

One important challenge for frequency-building approaches is that building up fluency in basic skills does not necessarily lead to the automatic transfer of knowledge in applied settings (Kubina and Yurich, 2012). Furthermore, the ability to apply critical thinking skills learnt in real-world or subject-specific contexts does not often come intuitively (Paul & Elder, 2009). One way to address these challenges is to use frequency building synergistically with instructional approaches that promote the transfer and the application of critical thinking skills across domains. For example, embedding critical thinking training into content-focused courses or instructions (Braun, 2004; Gray, 1993; Ikuenobe, 2001) can facilitate the transfer of critical thinking skills by teaching students 'how to think' rather than 'what to think' (Clement, 1979). Similarly, Halpern (1998) proposed a model for the trans-contextual learning of critical thinking skills, which scaffolds the learner's ability to apply skills in real-world contexts.

Current Study

In this study, we evaluated the effectiveness of an online learning intervention that aimed to enhance the critical-thinking skills of university students. The intervention focused on the ability of students to identify a type of reasoning error referred to as informal logical

fallacies (Carey, 2000). This ability is thought of as a hallmark component of critical thinking (Carey, 2000; Ramasamy, 2011).

The intervention adopted a bite-sized video-learning approach and used frequency building within a precision-teaching framework. We compared the learning performance of three experimental groups: a PT intervention group, a PT+ intervention group, and a self-directed learning control group. The two intervention groups (PT & PT+) received frequency-building practice aimed at increasing the rate of fallacy identification, with the addition of problem-based training in the PT+ group. The control group was exposed to the same instructional materials as the intervention groups but was asked to navigate through it in a self-paced way.

We examined students' learning of the taught critical thinking skills, as well as their ability to transfer taught knowledge in novel settings and to retain knowledge. More specifically, we measured student performance on the testing material in which they received instruction, as well as their performance in unseen examples and domain-general assessments of broader fallacy-identification skills. Follow-up assessments were repeated one week after the intervention. With these measures, we aimed to address the following research questions:

RQ1: Are there any educational benefits of frequency-building practice on students' learning of taught critical thinking materials?

RQ2: Are there any educational benefits of frequency-building practice on students' abilities to apply critical-thinking skills in novel contexts?

RQ3: How does frequency building affect students' knowledge retention following the intervention?

RQ4: Does the combination of frequency building with problem-based training support further benefits in students' learning of taught critical thinking materials (RQ1), generalisation in novel contexts (RQ2) or knowledge retention (RQ3)?

Instructional Framework for Teaching Critical Thinking Skills

Traditionally, critical-thinking training follows either the domain-general or the domain-specific approach (Tiruneh, Cook, & Elen, 2018). However, here, and in-line with other researchers (e.g. Koslowski, 1996; McNeill & Krajcik, 2009; Tiruneh et al., 2018), we take the view that domain-general and domain-specific expertise do not develop in isolation. Rather, both domain-general and context-specific knowledge is important for the effective acquisition of critical-thinking skills (McNeill & Krajcik, 2009). Thus, our instructional framework combines domain-general and domain-specific approaches. Specifically, the introduction to fallacy identification within bite-sized videos and frequency-building practice drew on elements of the domain-general approach; as learners could apply the critical-thinking skills learned across different domains. Whereas, problem-based training drew on elements of the domain-specific approach; as learners could learn how the skills are applied within subject-specific domains.

The domain-general approach is based on the assumption that the identification of informal logical fallacies shares commonalities across disciplines, and proficiency in this skill could transcend across the domain in which training was done. For example, lets consider a hypothetical Argument 1 “there is no proof that the parapsychology experiments were fraudulent, so I’m sure they weren’t.” and another hypothetical Argument 2 “because scientists cannot prove that global warming will occur, it probably won’t.”. Although the two arguments differ in terms of context (the first case involves a psychology matter, the second case involves a scientific discussion), both arguments are fallacious and share commonalities of using the lack of evidence as a proof of correctness (i.e. appeal to ignorance fallacy). In this study, scaffoldings of generic fallacy-identification skills within the bite-sized videos help students develop the ability to identify arguments that are “*psychologically* persuasive but *logically* incorrect” (Copy & Burgess-Jackson, 1996, p. 97). The exposure to structural features of

fallacies and the use of real-world examples within frequency-building practice prompt students to apply generatively what they had learned. This strategy aligns with Engle and colleagues' (2003) suggestion for *intercontextuality* as a means of bridging the gap between learning and transfer practices.

In addition, and following the domain-specific view, we also assume that critical-thinking skills may require explicit instruction within subject-specific domains to perform competently. This notion is similar to the Infusion approach, which emphasises how a critical-thinking skill could be applied within a subject-specific context (Abrami et al., 2008). In this study, the context-based scaffolding (i.e. problem-based training) within the PT+ group prompts students to apply critical-thinking skills in a context-specific situation. While we compare critical-thinking abilities between students in the PT and the PT+ groups, we, therefore, investigated if Infusion is necessary to promote the development of critical-thinking skills across domains (RQ4).

Method

Participants

A total of 57 adults (39 females, 17 males, 1 preferring not to say) with a mean age of 24.14 years (*SD* = 5.62; range 18-47 years old) took part in this study. Participants were recruited through the University's Research Participation System and departmental social media platforms. All participants were university students, with 37 registered as undergraduate students and 20 as postgraduate students. The study was approved by the Research Ethics Committee of the Department of Psychology.

Material

The intervention focused on four informal logical fallacies: 'appeal to ignorance', 'bandwagon', 'false cause', and 'hasty generalisation'. These four logical fallacies corresponded to common reasoning errors and were selected after consultation with a subject matter expert

(a senior lecturer of a university-level course involving critical thinking) and reviews of relevant textbooks (e.g., Gray, 1991; Schick & Vaughn, 2020). The four logical fallacies share a similar form, consisting of a premise followed by a conclusion (Fox & Ghezzi, 2003; see Table 1).

[Insert Table 1 around here]

Instructional Material

Learning Videos

Two 'bite-sized' learning videos, lasting 2:46 and 2:54 minutes, were created using the video animation software, Powtoon (https://www.powtoon.com). Powtoon has been highlighted as user-friendly software for supporting digital-based learning as it is equipped with various functions that can help to improve teacher's creativity, boost learning motivation, and support the learning needs of students with different abilities (Muhammad Basri et al., 2021; Resmol & Leasa, 2022; Zamora, Bravo, & Padilla, 2021).

Within the two learning videos, the first video (Episode 1: *Arguments and Fallacies*) presented learners with the standard form of an argument and introduced the four fallacies. The second video (Episode 2: *Examples of Fallacies*) gave examples for each of the four fallacies and explained why the arguments involved were fallacious or problematic.

Learning Tasks

Two learning tasks (one for each episode) consisting of 20 multiple-choice items were developed to facilitate knowledge acquisition after the presentation of the learning videos. Items for these tasks were based on material from the critical thinking textbooks (Gray, 1991; Schick & Vaughn, 2020) and were also reviewed by the subject-matter expert. Each item presented participants with a short paragraph that illustrated an example or a definition of a fallacy, followed by a forced-choice question asking participants to identify the relevant

fallacy. Participants received programmed feedback (“Correct!” or “Incorrect!”) on the screen after each answer selection.

Problem-based Tasks (used in the PT+ intervention group only)

Three problem-based tasks were developed to support learning in the PT+ intervention group, following each learning episode. The problem-based tasks consisted of open-ended questions, which required participants to analyse, evaluate, and explain flaws in reasoning within a psychological debate or dispute. Each task first presented a debate situation. This was done by showing a newsletter article or a short paragraph which summarised research findings referring to the main claim in dispute, alongside some context about the debate. For example, participants were presented with a paragraph entitled "does social media do more harm than good?" and referring to a recent survey, which found that feelings of loneliness among young workers increased as they reported higher amounts of time spent on social media. Then, participants were invited to identify fallacies in arguments presented by three panel members, who advocated for the disadvantages of social media (open-ended question, "Review the reasoning of each of the panel members A, B, and C and explain what might be problematic with their reasoning if considered to be faulty"). For example, a panel member would suggest that social media is doing more than good based on the fact that too much social media use will cause someone to feel lonely ('false cause'), and his friend, George, who uses social media more than 16 hours a day has been diagnosed to have depression lately ('hasty generalisation'). Participants were asked to review each argument and explain if a fallacy was involved.

Subsequently, participants were asked to indicate which of the three arguments presented by panel members they would be least likely to support (forced-choice question, “Indicate which one you believe to be the reasoning that you would be least likely to support”). Finally, participants were asked to provide a suggestion for the best course of action or the best counter-argument to resolve the debate (open-ended question, “If you are asked to give an

opinion in this debate, what would be your next course of action"). Programmed feedback was provided for each task following participants' responses to the questions involved. For example, the panel member above argued that there is a cause and effect relationship based on the correlation found and drew about the impacts of social media on all individuals on the basis of evidence concerning only certain people. Hence, the fallacies of false cause and hasty generalisation were committed.

Testing Material

Pre- and Post-Episode Tests based on the Learning Material

The questions included in the learning tasks of the two episodes were also used in the episode-specific tests of critical thinking. These were administered twice, at the beginning and the end of the episode. The pre- and post-episode tests were administered as time-based assessments (to consider both accuracy and speed in identifying the fallacies). Participants were instructed to answer the questions as accurately and as fast as they could within a minute. No feedback was given in the pre- and post-episode tests.

Pre- and Post-Intervention Assessments on Unseen Questions

An additional 50 multiple-choice questions were used to assess participants' ability to recognise fallacies in unseen questions. These were selected from the same bank of questions used for the development of the learning tasks and the pre-and post-episode tests. 25 items were presented as a pre-intervention assessment and the rest as a post-intervention assessment.

Broader Abilities in Fallacy Identification: Informal Reasoning Fallacies Identification Task (IRFIT; Neuman, 2003)

To assess the students' broader abilities in fallacy identification, we used a test based on the Informal Reasoning Fallacies Identification Task (Neuman & Weizman, 2003; Weinstock, Neuman, & Tabak, 2004). In this study, four informal reasoning tasks, each consisting of two items adapted from Neuman (2003)'s study, were administered to

participants. Each reasoning task corresponded to one of the four fallacies and consisted of an argumentative scenario followed by four questions. The scenario was structured in four sentences as follows. The first sentence presented participants with two debaters who were described as either psychology students or philosophers. The second sentence presented the context and the main claim under debate stated in the form of a question. The third and the fourth sentences presented the arguments by the two debaters, a so-called "protagonist" and an "antagonist". Finally, the specific reasoning of one of the debaters in support of their position was presented with a fallacy involved.

Participants were asked to identify potential flaws in reasoning and identify fallacies. In particular, they responded to the following four questions:

1) A *yes/no fallacy identification question*, which examined whether participants conceived an argument as fallacious or problematic (e.g. "Do you think there is a problem in the argument that the antagonist presented in Line 5?").

2) A *open-ended fallacy explanation question*, which assessed participants' ability to articulate what they perceived to be faulty with the reasoning of an argument (e.g. "If you think that there is a problem in the argument presented by the antagonist, what is the problem?").

3) An *open-ended response question*, which assessed participants' ability to debate and present a counter-argument (e.g. "What is the best answer the protagonist can use in response to the antagonist's argument?").

4) A *forced-choice fallacy classification question*, which assessed whether participants perceived the argument to be a quarrel, a formal debate, or a critical discussion (e.g. "In your opinion, what is the main reason for the debate between the two arguers"). Participants responded to this question by selecting one of the three answer choices: (a) They do not like each other and, therefore, each person is attacking the other's claim - quarrel, (b) Each one of

them wants to impress his colleagues and win the debate - formal debate, and (c) They have different opinions on this matter, and they are trying to convince each other - critical discussion.

Design

The design of the study is shown in Figure 1. Participants were randomly allocated to three groups: (A) a 'precision teaching (PT)' intervention group, (B) a 'precision teaching plus problem-based training (PT+)' intervention group, and (C) a 'self-directed learning' control group. The three groups were exposed to the same instructional material and testing stimuli; however, this was administered in different ways to implement different learning conditions. In particular, the PT group received frequency-building learning tasks, which aimed at increasing the rate of fallacy identification. The PT+ group completed frequency-building learning tasks combined with the addition of problem-based training to facilitate a better application of critical thinking in the PT condition. Finally, the control group completed learning tasks in a self-directed way.

[Insert Figure 1 around here]

Procedure

Participants completed the study in three sessions administered online via the Qualtrics platform (Qualtrics, Provo, UT). In the first online session, participants completed the pre-intervention assessment and Episode 1, *Arguments and Fallacies*. In the second online session, participants completed Episode 2, *Examples of Fallacies* and the post-intervention assessment. In the last online session, which was administered a week after the completion of Session 2, participants repeated both the post-episode assessments for Episode 1 and Episode 2 as retention assessments.

Each episode started with a time-based pre-episode assessment on fallacy identification. The assessment was followed by the participants watching a learning video, in which the definitions (Episode 1) or examples of fallacies (Episode 2) were explained for

approximately three minutes. Participants were asked to watch the video until the end, and the next button to proceed with the next part was only presented at the bottom of the page towards the end of the video presentation. Then, participants completed two blocks of 20 multiple-choice questions which were administered to the three groups as learning tasks in different forms. The learning tasks allowed participants to familiarise themselves with and consolidate knowledge learnt from the video content. Finally, participants completed the post-episode assessment within a 1-minute timeframe.

The three groups were differentiated in the types of learning tasks they completed within the two learning episodes, as detailed in the following section.

Learning tasks in the PT intervention group. Learning tasks in the PT intervention group were guided by a high response-rate requirement implemented in iterations of timed sprints and feedback. Participants were informed that they were going to practice identifying the fallacies within a 1-minute timeframe, with the remaining time appearing on the top left corner of the screen. They selected the best answer out of the four choices as fast as they could and received programmed feedback after each response ("Correct!" or "Incorrect!"). After the 1-minute timed sprint, participants were shown the number of accurate responses they had provided. Then, participants proceeded to an error-correction procedure, which focused on the questions they had answered incorrectly. During the error correction procedure, participants were instructed to answer these questions again, without any time limit, and were shown the accurate answer if they gave an incorrect response for a second time. After the error correction procedure, participants answered the 20 multiple-choice questions with the same procedure as the first timed sprint again. The error-correction procedure and the learning cycle were repeated twice before progressing to complete the post-episode test (see Figure 2).

[Insert Figure 2 around here]

Learning tasks in the PT+ intervention group. Participants who were assigned to the PT+ intervention group completed the same learning tasks as the PT group. Additionally, participants in this group completed the corresponding problem-based task following each episode.

Learning tasks in the control group. In this group, learning tasks were completed in a self-directed way, without a high response-rate requirement. Participants were instructed to answer all 20 questions accurately and as fast as they could (but not within timed sprints) and were given feedback on the number of correct responses they achieved. This cycle was repeated twice before progressing to complete the post-episode test. Hence, the main difference between the intervention groups (PT & PT+) and the control group was that participants in the control group did not complete the learning tasks in 1-minute timed sprints; rather, they were asked to complete the whole tasks at their own pace. The learning tasks and the number of blocks conducted in each episode remained the same as in the intervention groups.

Scoring

IRFIT

Content analysis was conducted on participants' answers to the tasks by two researchers. Using the scoring procedures from Neuman (2003)'s study, 10% of the data was marked by both scorers and Cohen's Kappa showed that there was strong agreement between the two scorers ($\kappa = .814$; McHugh, 2012). The yes/no identification question was scored as 1 for a 'yes' answer and 0 for a 'no' answer. Both open-ended fallacy explanation and response questions were marked as 1 when participants took into account to identify and/or explain the informal reasoning fallacy involved in the situation. Participants scored 0.5 when they captured the key elements of why the arguments were fallacious but nonetheless did not provide a complete explanation. Participants scored 0 when either they did not answer the question, did

not identify the problem in the situation, or did not take into account the fallacy involved when explaining.

Data Analysis

Quantitative data collected from the pre-and post-episode tests and the pre-and post-intervention assessments were analysed to examine the effects of time (within-participants factor) and differences between groups (between-participants factor). When preliminary data checks suggested that the assumptions of normality and homogeneity of variance were met, data were analysed with a 3 (Groups: PT vs. PT+ vs. control) x 2 (Time: pre- vs. post-episode/intervention) mixed-design ANOVA. When these assumptions were violated, Wilcoxon Signed Rank non-parametric tests (within-participants) were used to compare differences in a given measure across two time points, and Kruskal Wallis non-parametric tests (between-participants) were used to examine differences in the changes in the measure between groups. If the data were normal but the homogeneity of variance was violated, changes on a measure over time were examined with t-tests, and between-group differences in change over time were examined with a Welsch one-way ANOVA.

In a complementary analysis, we compared changes between participants with relatively low and relatively high performance.

In all analyses, effect sizes were reported using relevant standardised measures (t-tests: Cohen's *d*; Wilcoxon Signed Rank/Kruskal Wallis: *r*, Welch one-way ANOVA: ω^2, mixed ANOVA: ηp^2). For Cohen's *d* and *r*, a value of .20 was taken to suggest a small effect size, a ±.50 a medium effect size, and ±.80 a large effect size; for ω^2 and ηp^2 the thresholds were 0.01 (small), 0.06 (medium) and 0.13 (large) (Cohen, 1998). Effect sizes *d* greater than 0.40 were considered educationally relevant (Hattie, 2009).

Results

Pre- and post-episode tests on the learning tasks

Figure 3 presents the mean scores of the pre- and post-episode tests for Episode 1 and 2 for the three groups. Shapiro-Wilk tests indicated that the assumption of normality was not met ($p < .05$ for Episode 1 pre- and post-episode tests, and Episode 2 pre-episode test), hence, Wilcoxon Signed Rank tests were conducted to examine the changes in performance within each episode. The results showed that participants on average, scored significantly higher in the post-episode (Episode 1: *Mdn* = 10.00; Episode 2: *Mdn* = 10.00) compared to the pre-episode tests (Episode 1: *Mdn* = 5.00; Episode 2: *Mdn* = 5.00) on the learning tasks, for both Episode 1 ($Z = 6.31$, $p < .001$, $r = 0.84$) and Episode 2 ($Z = 5.78$, $p < .001$, $r = .77$).

Given that the data were not normally distributed, we compared improvements in the three groups using Kruskal Wallis tests for Episode 1 (PT: $Mdn_{\text{difference}} = 5.00$; PT+: $Mdn_{\text{difference}} = 6.00$; Control: $Mdn_{\text{difference}} = 5.00$) and Episode 2 (PT: $Mdn_{\text{difference}} = 6.00$; PT+: $Mdn_{\text{difference}} = 5.00$; Control: $Mdn_{\text{difference}} = 4.00$). These tests suggested that the improvements of the three groups were comparable in both Episode 1 [$H(2) = 0.17$, $p = 0.920$, $r = 0.02$] and Episode 2 [$H(2) = 1.02$, $p = 0.601$, $r = 0.13$].

[Insert Figure 3 around here]

Pre- and post-intervention assessments on unseen questions

Figure 4 shows mean accuracy scores in the pre-and post-intervention for the three groups. Shapiro-Wilk tests indicated that all data were statistically normal (all *ps* > .05). However, the preliminary Levene's test suggested that the assumption of homogeneity of variance was not met for the post-test measures ($p = .017$).

Paired sample t-tests were thus conducted to compare performance between pre-and post-intervention assessments in the three groups. These tests suggested significant

improvements in all three groups [PT: t (18) = 10.33, p < .001, d = 2.37; PT+: t (18) = 7.68, p < .001, d = 1.76; Control: t (18) = -4.12, p = .001, d = 0.95].

To compare participants' improvements between groups, a Welch one-way ANOVA with corrected degrees of freedom was used. The results showed a trend for a difference between the average scores of the three groups, which, however, did not reach levels of statistical significance, F (2, 34.63) = 2.61, p = .088, ω^2 = 0.05.

[Insert Figure 4 around here]

To gain further insight into the non-significant trend of between-group differences, in a complementary analysis, we divided participants into lower- and higher-scorer categories based on their pre-test scores. Participants who scored at the 50th percentile and below were categorised as lower-scorers (n = 33), and those who scored above the 50th percentile were categorised as higher-scorers (n = 24). Figure 5 shows mean accuracy scores of low- and high-scoring participants in the pre-and post-test. Shapiro-Wilk tests indicated that the assumption of normality was not met for the pre-and post-test scores (all ps < .05). Hence, a Wilcoxon Signed Rank non-parametric test was conducted to compare participants' scores between pre- and post-intervention assessments. The results showed that both low- and high-scoring participants achieved significantly higher mean scores at post-intervention compared to pre-intervention (Low-scoring: Z = 4.79, p < .001, r = 0.83; High-scoring: Z = 3.68, p < .001, r = 0.75].

With regards to differences in the improvement of low- and high-scoring participants, a Kruskal Wallis test suggested a significant difference, H (1) = 4.48, p = 0.034, r = 0.59, with larger improvements for low-scoring ($Mdn_{\text{difference}}$ = 6.00) than for high-scoring participants ($Mdn_{\text{difference}}$ = 4.50).

[Insert Figure 5 around here]

Pre- and post-intervention assessment on broader critical thinking skills (IRFITs)

Figure 6 shows the average scores of the three groups in the IRFIT, the assessment of how well participants applied their critical thinking skills in a broader context of fallacy identification. These data were analysed with parametric statistics; in particular, a 3 x 2 mixed-design ANOVA was conducted, with Group as a between-subjects factor and Time as a within-subjects factor. The analysis showed a significant main effect of Group, F (2, 54) = 6.09, p = .004, η_p^2 = .184 ('large' effect), which was further explored with posthoc comparisons. These suggested that the performance scores for the PT (M = 11.29) and the PT+ intervention groups (M = 12.20) were higher than the scores of the control group (M = 9.07) (PT vs. Control: p = .216; PT+ vs. Control: p = .007, PT vs PT+: p = .127). There was also a significant main effect of Time, F (1, 54) = 9.82, p < .003, η_p^2 = .154, whereby the post-intervention score (M = 11.35) was higher than the pre-intervention score (M = 10.35), as well as a significant interaction between the two factors, F (2, 54) = 4.14, p = .021, η_p^2 = .133 (see Figure 6), reflecting a significant improvement for the PT (p = 0.001) and PT+ (p = 0.046) intervention groups but not the control group.

[Insert Figure 6 around here]

Knowledge retention

Figure 7 shows the average scores of the three groups in the post-episode assessments and the retention tests for Episode 1 and Episode 2. Shapiro-Wilk tests indicated that all data were statistically normal. Levene's tests also showed that the assumption of homogeneity of variance was met. Thus, the data were analysed with a 3 x 2 mixed-design ANOVA with Group as a between-subject factor, and Time (post-episode assessment vs. retention test) as a within-subject factor. For Episode 1, the results showed no significant main effect of Group, F (2, 48) = 0.22, p = .803, η_p^2 = .007; Time, F (1, 48) = 3.00, p = .090, η_p^2 = .011; and no interaction, F (2, 48) = 1.35, p = .269, η_p^2 = .009. Similarly, for Episode 2, there was no significant main

effect of Group, F (2, 47) = 0.71, p = .497, η_p^2 = .023; Time, F (1, 47) = 3.26, p = .077, η_p^2 = .015; and no interaction, F (2, 47) = 0.40, p = .676, η_p^2 = .004.

[Insert Figure 7 around here]

Discussion

In this study, we implemented and evaluated an online learning design aiming to improve critical thinking skills in university students based on a video-based learning approach that used frequency building under precision teaching. We also combined the frequency-building approach with structured problem-based training to further foster the transfer of the taught skills. We compared the learning performance of the three experimental groups, examining students' performance in the taught materials, in unseen examples, and in more general fallacy-identification problems, as well as in follow-up tests.

With regards to whether PT could improve students' learning of the taught material (RQ1), our results from the post-episode tests demonstrated that all groups showed significant and comparable improvements in their ability to identify the taught examples of fallacies. Thus, all three types of learning condition, PT-based and not, worked equally well in supporting video-based teaching of fallacy-identification and yielded comparable outcomes, in line with findings from an earlier study from Fox and Ghezzi (2003)'s. Furthermore, taking into account that the broader PT literature tends to focus on simpler and low-level skills, our current findings are important because they suggest that the use of precision teaching can be extended to complex and high-level skills such as critical thinking (Cuzzocrea et al., 2011).

With regards to the application of the taught knowledge into unseen examples (RQ2), the analyses of the post-intervention assessments suggested that, again, all learning conditions yielded comparable improvements. Interestingly, these improvements were greater for students who scored at or below the 50th percentile. Although this result could be, partially, attributed

to a ceiling effect, it demonstrates the usefulness of technology-enhanced learning designs to all students and especially those who present difficulties in critical thinking.

Turning to the transfer of the taught skills in a domain-general IRFIT task (RQ2), our results showed that, importantly, only the two PT groups showed reliable improvements in performance post-intervention. Thus, frequency building under the precision teaching framework can serve to foster the application of skills in novel contexts, in line with Kubina and Yurich (2012), who suggested that frequency-building can lead to desirable outcomes of knowledge generalisation.

Furthermore, in the knowledge retention tests (RQ3), there were no significant differences between the post-episode assessment scores and the retention test scores, implying that students, regardless of groups, presented non-significant detriments in their fluency even after a week without practice. Earlier research suggested that the frequency-building practice can support the retention of skills for a longer period of time (Binder, 1996). However, as the difference between the intervention and the control groups was not significant in our study. To understand this inconsistency between our findings and earlier research, further investigation into how frequency-building procedures impact long-term retention is warranted, possibly by extending the time point of retention tests.

With regards to whether problem-based training can support further benefits in students' acquisition, generalisation, and retention of critical-thinking skills (RQ4), improvements in the domain-general task learning were comparable in the PT and PT+ group, suggesting that problem-based training is, indeed, not necessary for promoting the transfer of taught skills. This finding is in contrast with previous literature positing that rigorous practice for critical thinking is required until students can internalise the concepts learnt and demonstrate critical thinking skills intuitively in their daily lives (Paul & Elder, 2009).

In sum, the current study provides a foundation for understanding how the use of video technologies and frequency-building practice can be combined into an effective supplementary teaching tool to promote critical thinking in online settings. The integration of the two approaches is suitable for supporting students of various abilities. Our instruction framework draws on elements from Papert's constructivism, in which effective learning occurs by building upon individual students' prior knowledge through active engagement (Papert, 1980). In this study, the use of video technologies to present learning information in a "bite-sized" format helps to maximise students' engagement with the content and offers students the flexibility to pause, rewind, and revisit any part of the video whenever necessary (Saline et al., 2012). The inclusion of online frequency-building intervention also improves the quality of the session, as it transforms it from solely a passive video-watching event to an active learning opportunity that helps students monitor their own learning and is necessary for knowledge construction (Gaudin & Chaliès, 2015).

This online learning approach addresses challenges in critical thinking instructional designs related to promoting active learning during students' independent study time (Mandernach, 2006). Our study shows that this type of practice, which focuses on building fluency of skills, is flexible enough to be used in teaching complex concepts such as critical thinking and could lead to desirable learning outcomes, specifically, on the application of skills in a novel setting. Moreover, our study demonstrated that the online learning design of frequency-building intervention is accommodative to individual students, offering students the opportunity to practice their individual mistakes following each practice trial. A technology-enhanced model of frequency-building practice like this also allows a systematic presentation of stimuli and effective tracking of student engagement (Beverley, Hughes, & Hastings, 2009). Our approach to teaching critical thinking skills is versatile and also applicable to the current

landscape in Higher Education which the COVID-19 restrictions have transformed (Pokhrel & Chhetri, 2021).

Limitations directions for further research

Our study is not without limitations. First and in terms of scope, our intervention focused on fallacy identification. However, critical thinking is a multifaceted construct, and future studies should be inclusive of more diverse processes related to the interpretation, analysis, evaluation, and inference, such as argument analysis, evaluation of the credibility of claims or sources, and identification of scientific versus pseudo-scientific procedures (McPeck, 1981).

Furthermore, in terms of research methodology, although participants in the three groups were exposed to similar instructional materials and procedures, the time of exposure in the learning task was not controlled. A more nuanced investigation of learning under precision teaching will need to explicitly examine the engagement with the learning materials. This is important as it has been argued that frequency-building procedures can reduce the time needed to master a targeted skill (Lokke et al., 2008). Furthermore, in the current study, a short-duration precision-teaching intervention yielded significant improvements in fallacy identification performance in novel problem-solving contexts – albeit a small one.

An additional limitation lies in the use of random group allocation in our experimental design, rather than controlling for the participants' demographics across experimental groups. In this study, participants were randomly allocated to three groups that were exposed to the same instructional stimuli but differed in the way that the learning tasks were performed. Random allocation has been widely used in educational research to evaluate the effectiveness of interventions and to ensure that any group differences are due to chance (Forsetlund et al., 2007). Nevertheless, we acknowledge that there might be individual variations in participants' educational level, enrolled course, and motivation to learn that we did not account for in this

study. One could draw more robust conclusions by assessing how the impact of this intervention depended on these demographics.

Finally, in this study, we did not include instructors in the learning videos. Instead, we used animated videos created using the Powtoon platform. This decision was partly influenced by the time when the research was developed. COVID-19 lockdown restrictions were in place, and all physical engagements were halted during that period, limiting our ability to carry out video recordings with an instructor in place. While various studies have highlighted the benefits of Powtoon-based videos on student engagement and motivation (Muhammad Basri et al., 2021; Zamora, Bravo, & Padilla, 2021), contrasting evidence suggests that some students find learning videos featuring a presenter to be more engaging (Guo et al., 2014; Pi, Hong, & Yang, 2017). Future studies could examine the impact of the presence of instructors on students' engagement and learning. An interesting possibility is to consider peers as presenters as evidence suggested that perceived similarity between a peer and the learner could create a favourable learning environment that can benefit learning (Bulte, Betts, Garner, & Durning, 2007; Lockspeiser, O'Sullivan, Teherani, & Muller, 2008).

Conclusion

The current study demonstrated the potential of an online intervention approach of video-based learning and PT to improve critical-thinking skills of university students. After a brief intervention, which consisted of only two learning episodes, students showed improvements in fallacy identification performance, which transferred into novel problem-solving contexts. These results are important in an era of over specialisation, in which critical thinking is identified as one of the most desired yet most challenging educational outcomes for Higher Education. Given the increased use and acceptance of technology-enhanced approaches as a result of the recent transformation of the Higher Education landscape following the

COVID-19 restrictions, the current results provide a new perspective for the combination of video learning and PT practice in an online learning environment.

Figures and Tables

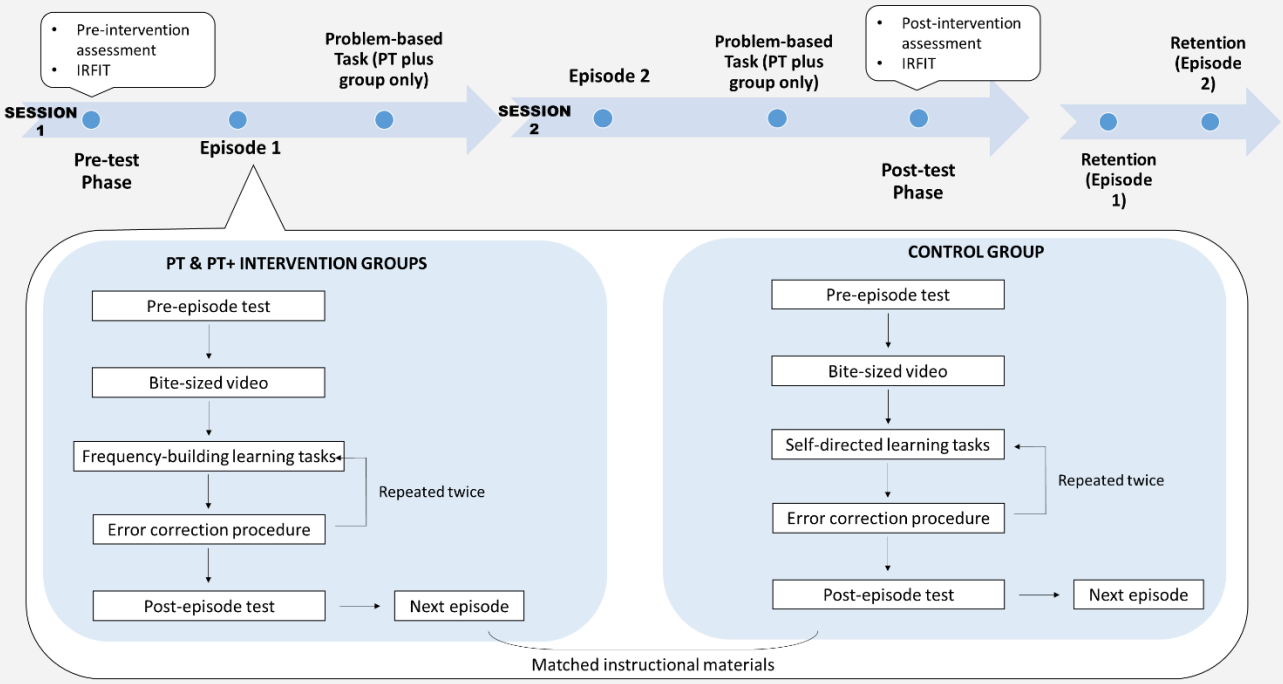
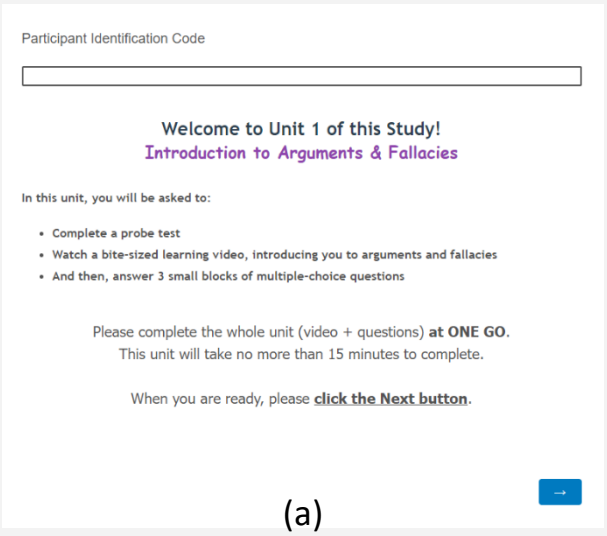
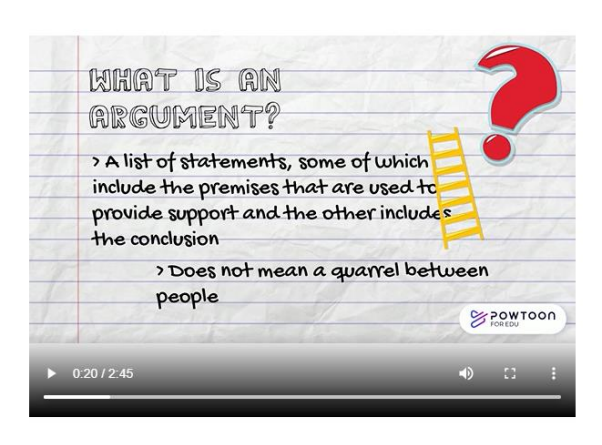


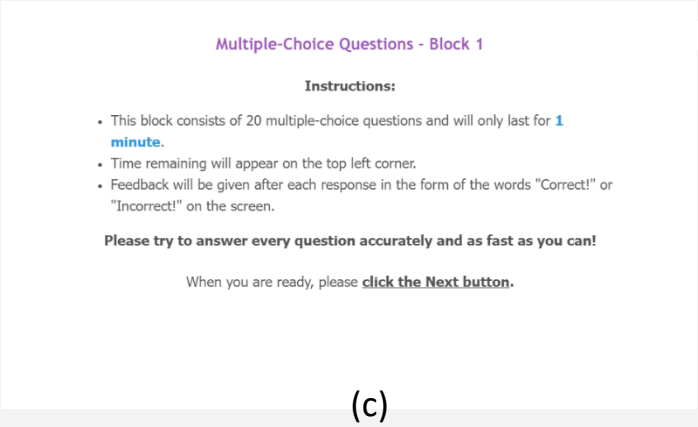
Figure 1. *Flowchart of the Study*.



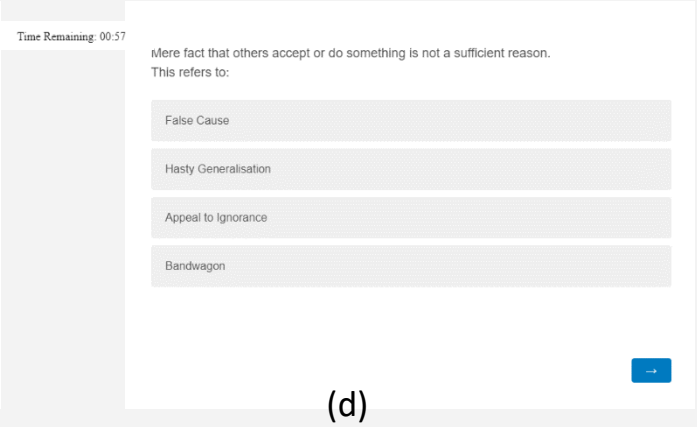
(a)



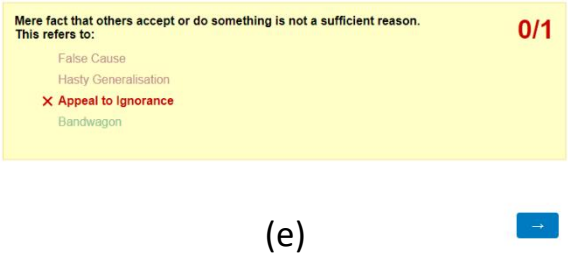
(b)



(c)



(d)



(e)

Figure 2. *Screenshots of the learning tasks interface for PT intervention groups – (a) instruction page; (b) video presentation page; (c) block presentation page; (d) learning tasks page; (e) error correction procedure page*.

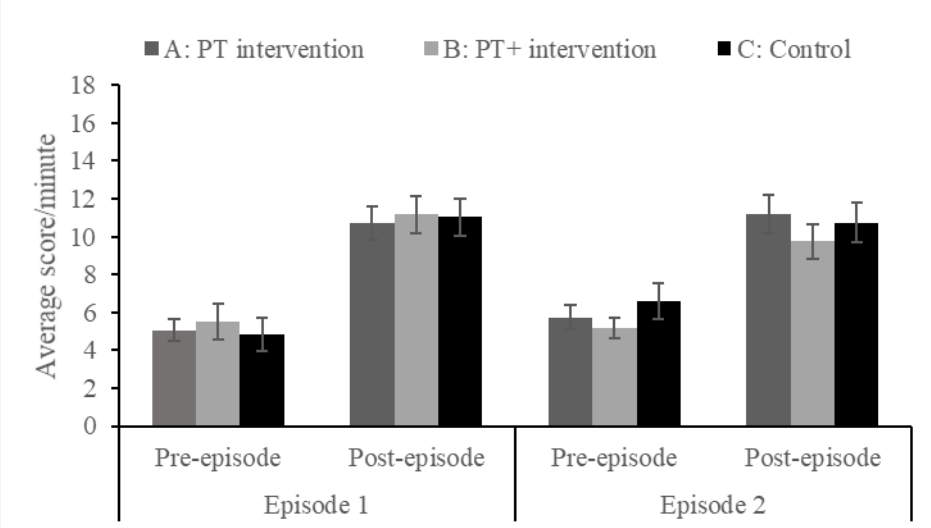


Figure ~~2~~3. *Mean scores of the pre-and post-episode tests. Scores were calculated out of participants' accurate responses to 20 questions within a minute. Error bars represent standard errors of the means.*

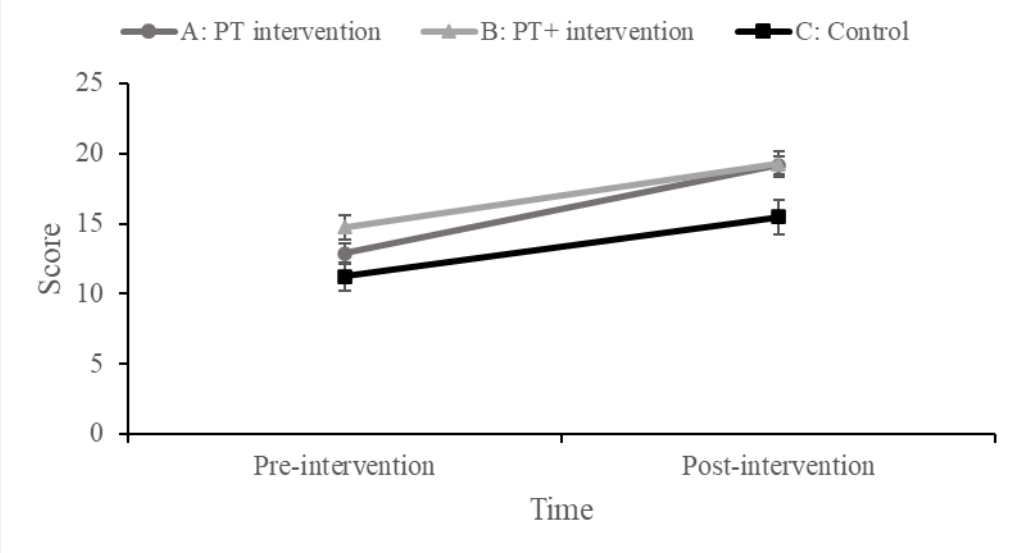


Figure ~~3~~4. *Mean accuracy scores of the pre-and post-intervention assessments. Scores were calculated out of 25 questions. Error bars represent standard errors of the means.*

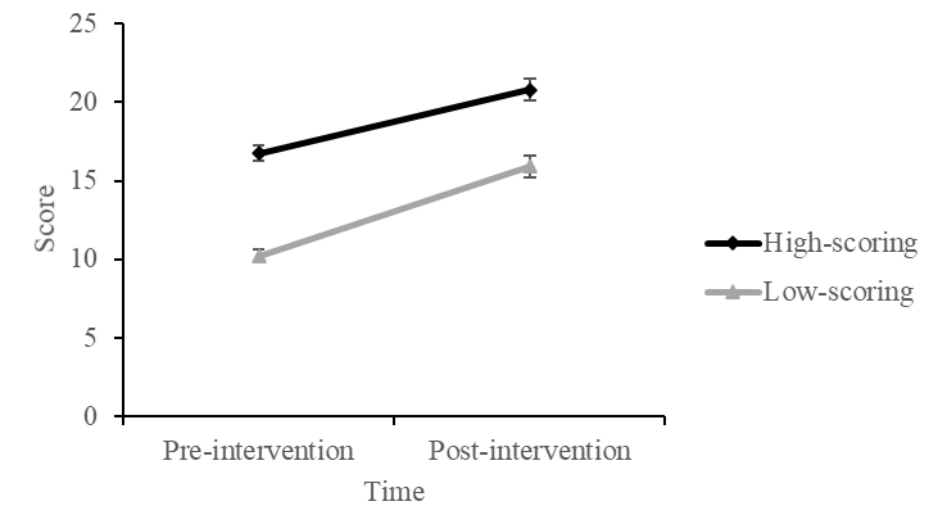


Figure 45. *Mean accuracy scores for low- and high-scoring participants at pre-and post-intervention assessments. Scores were calculated out of 25 questions. Error bars represent standard errors of the means.*

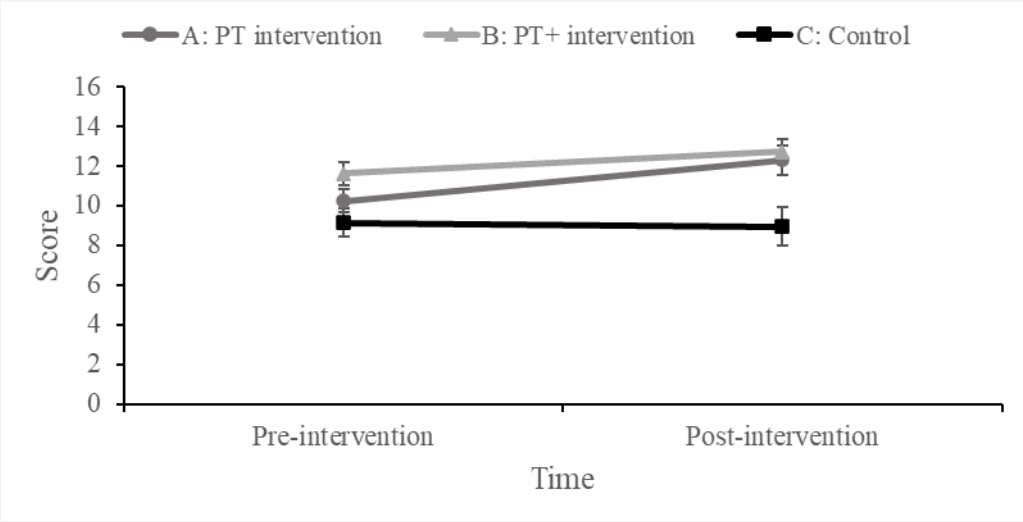


Figure ~~5~~6. *Mean performance scores of the IRFITs at pre-and post-intervention. Scores were calculated out of four IRFITs at each time point. Error bars represent standard errors of the means.*

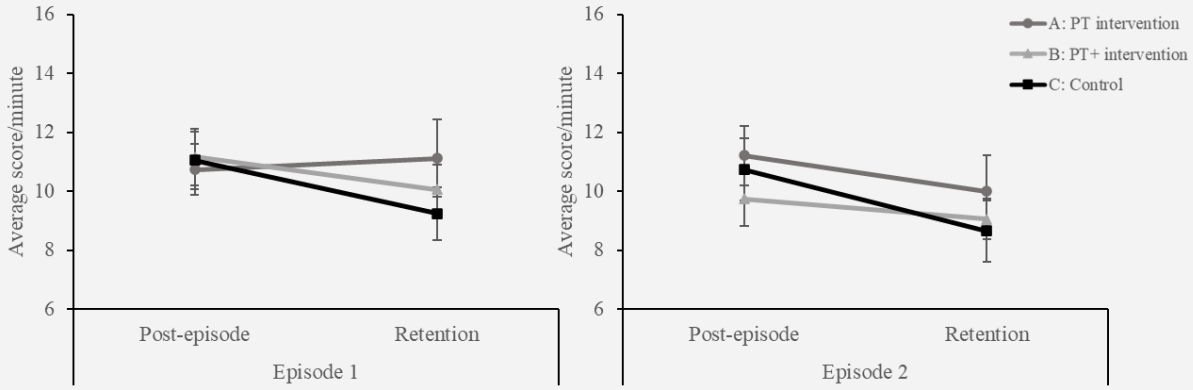


Figure ~~6~~7. *Mean scores for all three groups at the post-episode assessments and the retention tests of Episode 1 and Episode 2. Scores were calculated out of participants' accurate responses to 20 questions within a minute. Error bars represent standard errors of the means.*

Table 1

Informal Logical Fallacies

Fallacy	Definition	Common Form	
Appeal to Ignorance	Arguing that something is true because it has not been proved false	Premise	X is not known to be false.
		Conclusion	Therefore, X is true.
Bandwagon	Arguing that something must be true or good because it is in accordance with the beliefs of the majority	Premise	Everybody believes that X is true.
		Conclusion	Therefore, X is true.
False Cause	Assuming a causal link between a premise and a conclusion based on an imagined causal connection that probably does not exist	Premise	X correlates with Y.
		Conclusion	Therefore, X causes Y.
Hasty Generalisation	Drawing a general conclusion about all elements of a certain type/category based on partial evidence, from some elements of that type/category	Premise	X has been shown to be true in some instances.
		Conclusion	Therefore, X is true in all instances.

References

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. doi: 10.3102/0034654314551063

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M., 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. doi:10.3102/0034654308326084

Baker, D. C. (2001). Nursing reasoning model. *Nurse Educator, 26*(5), 203-204.

Beverley, M., Hughes, J. C., & Hastings, R. P. (2009). What's the probability of that? Using SAFMEDS to increase undergraduate success with statistical concepts. *European Journal of Behaviour Analysis, 10*(2), 235-247. doi: 10.1080/15021149.2009.11434321

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19*(2), 163-197.

Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximising student learning from video content. *CBE—Life Sciences Education, 15*(4), es6.

Braun, N. M. (2004). Critical thinking in the business curriculum. *Journal of Education for Business, 79*(4), 232–236.

Broadbear, J. T. (2003). Essential elements of lessons designed to promote critical thinking. *Journal of the Scholarship of Teaching and Learning,* 3(3), 1-8.

Brunt, B. A. (2005). Models, measurement, and strategies in developing critical-thinking skills. *The journal of continuing education in nursing*, *36*(6), 255-262.

Bulte, C., Betts, A., Garner, K., & Durning, S. (2007). Student teaching: Views of student near peer teachers and learners. *Medical Teacher, 29*, 583-590. doi: 10.1080/01421590701583824

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.

Carey, S. S. (2000). *The uses and abuses of argument: Critical thinking and the fallacies.* Mountain View, CA: Mayfield.

Carmichael, M., Reid, A-K., & Karpicke, J. D. (2018). *Assessing the impact of educational video on student engagement, critical thinking and learning: The current state of play* [White Paper]. Retrieved from https://us.sagepub.com/sites/default/files/hevideolearning.pdf

Chiesa, M., & Robertson, A. (2000). Precision teaching and fluency training: Making maths easier for pupils and teachers. *Educational Psychology in Practice, 16*(3), 297–310. Retrieved from http://doi.org/10.1080/713666088

Clement, J. (1979). Introduction to research in cognitive process instruction. In Lochhead, J. and Clement, J. (Eds.), *Cognitive process instruction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

Commons, M. L., Owens, C. J., & Will, S. M. (2015). Using a computer-based precision teaching program to facilitate learning of complex material: The case of the model of hierarchical complexity. *Behavioral Development Bulletin, 20*(2), 207.

Copy, I. M., & Burgess-Jackson, K. (1996). *Informal logic.* NJ: Prentice Hall.

Cuzzocrea, F., Murdaca, A. M., & Oliva, P. (2011). Using Precision Teaching method to improve foreign language and cognitive skills in university students. *International Journal of Digital Literacy and Digital Competence (IJDLDC), 2*(4), 50-60. doi:10.4018/jdldc.2011100104

Engle, R., & Greeno, J. (2003, April). Framing interactions to foster productive learning. In K. Beach (Chair), *Sociocultural, semiotic, situative, and activity theoretic alternatives to the transfer metaphor: New understandings of how knowledge generalizes.* Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL

Facione, P. A., & Facione, N. C. (2001). Analysing explanations for seemingly irrational choices: Linking argument analysis and cognitive science. *International Journal of Applied Philosophy, 15*(2), 267–286. Retrieved from https://insightassessment.com/wp- content/uploads/ia/pdf/IJAP_Analysis_Paper.pdf

Forbes, K. (2018). Exploring first year undergraduate students' conceptualisations of critical thinking skills. *International Journal of Teaching and Learning in Higher Education*, *30*(3), 433-442.

Forsetlund, L., Chalmers, I., & Bjørndal, A. (2007). When was random allocation first used to generate comparison groups in experiments to assess the effects of social interventions?. *Economics of Innovation and New Technology*, *16*(5), 371-384.

Fox, E. J., & Ghezzi, P. M. (2003). Effects of computer-based fluency training on concept formation. *Journal of Behavioral Education, 12*(1), 1-21.

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review, 16*, 41–67. Retrieved from https://doi.org/10.1016/j.edure v.2015.06.001

Guo, P. J., Kim, J., & Rubin, R. (2014, March). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 41-50).

Gray, P. (1993). Engaging students' intellects: The immersion approach to critical thinking in psychology instruction. *Teaching of Psychology, 20*(2), 68-74.

Gray, W. D. (1991). *Thinking critically about New Age ideas.* Belmont, CA: Wadsworth Publishing Company.

Guiller, J., Durndell, A., & Ross, A. (2008). Peer interaction and critical thinking: Face-to-face or online discussion?. *Learning and instruction*, *18*(2), 187-200.

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

Harrington, K., Norton, L., Elander, J., Lusher, J., Aiyegbayo, O., Pitt, E., Robinson, H., & Reddy, P. (2006). Using core assessment criteria to improve essay writing. In Bryan, C. & Clegg, K. (Eds.), *Innovative assessment in higher education* (pp. 110-119). London: Routledge.

Hart Research Associates. (2015). *Falling short?: College learning and career success.* Washington, D.C.: Association of American Colleges and Universities (AAC&U). Retrieved from http://www.aacu.org/leap/public-opinion-research/2015-survey-results

Hatcher, D. L. (2011). Which test? Whose scores? Comparing standardised critical thinking tests. *New Directions for Institutional Research, 149*, 29-39. doi: 10.1002/ir.378

Hattie, J. (2009). *Visible learning.* Routledge.

Hill, J. L., & Nelson, A. (2011). New technology, new pedagogy? Employing video podcasts in learning and teaching about exotic ecosystems. *Environmental Education Research, 17*(3), 393-408. doi: 10.1080/13504622.2010.545873

Hughes, J. C., Beverley, M., & Whitehead, J. (2007). Using precision teaching to increase the fluency of word reading with problem readers. *European Journal of Behavior Analysis, 8*(2), 221-238.

Ikuenobe, P. (2001). Teaching and assessing critical thinking abilities as outcomes in an informal logic course. *Teaching in Higher Education, 6*(1), 19-32.

Joynes, C., Rossignoli, S., & Fenyiwa Amonoo-Kuofi, E. (2019). *21st Century Skills: Evidence of issues in definition, demand and delivery for development contexts (K4D Helpdesk Report)*. Brighton, UK: Institute of Development Studies.

Khong, H. K., & Kabilan, M. K. (2020). A theoretical model of micro-learning for second language instruction. *Computer Assisted Language Learning*, 1-24.

Koh, N.S., Gottipati, S., & Shankararaman, V. (2018). Effectiveness of bite-sized lecture on student learning outcomes. In *4th International Conference on Higher Education Advances (HEAd'18)*, 515-523. doi: http://dx.doi.org/10.4995/HEAd18.2018.8027

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning.* Cambridge, MA: MIT Press.

Kreth, M., Crawford, M. A., Taylor, M., & Brockman, E. (2010). Situated assessment: Limitations and promise. *Assessing Writing*, *15*(1), 40-59.

Kubina, R. M., & Morrison, R. S. (2000). Fluency in education. *Behavior and Social Issues, 10*, 83-99.

Kubina, R. M., & Yurich, K. K. (2012). *The precision teaching book.* Lemont, Pennsylvania: Greatness Achieved Publishing Company.

Lindsley, O. R. (1997). Precise instructional design: Guidelines from precision teaching. *Instructional development paradigms*, 537-554.

Lockspeiser, T. M., O'Sullivan, P., Teherani, A., & Muller, J. (2008). Understanding the experience of being taught by peers: The value of social and cognitive congruence. *Advances in Health Sciences Education, 13*(3), 361-372. Retrieved from https://doi.org/10.1007/s10459-006-9049-8

Lokke, G. E., Lokke, J. A., & Arntzen, E. (2008). Precision Teaching, Frequency-Building, and Ballet Dancing. *Journal of Precision Teaching and Celeration, 24*, 21-27.

MacKnight, C. B. (2000). Teaching critical thinking through online discussions. *Educause Quarterly, 23*(4), 38–41. Retrieved from http://eac595b.pbworks.com/f/macknight+2000+questions[1].pdf

Mandernach, B. J. (2006). Thinking critically about critical thinking: Integrating online tools to promote critical thinking. *Insight: A collection of faculty scholarship*, *1*, 41-50.

Marton, F. (2006). Sameness and difference in transfer. *The journal of the learning sciences*, *15*(4), 499-535.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica, 22*(3), 276-282.

McNeill, K. L., & Krajcik, J. (2009). Synergy between teacher practices and curricular scaffolds to support students in using domain-specific and domain general knowledge in writing arguments to explain phenomena. *The Journal of the Learning Sciences, 18*(3), 416-460. doi: 10.1080/10508400903013488

McPeck, J. (1981). *Critical thinking and education.* Toronto, Ontario, Canada: Oxford University Press.

Merrill, M. D., Tennyson, R. D., & Posey, L. O. (1992). *Teaching concepts: An instructional design guide* (2nd ed.). Englewood Cliffs, NJ: Educational Technology Publications.

Neuman, Y. (2003). Go ahead, prove that God does not exist! On high school students' ability to deal with fallacious arguments. *Learning and Instruction*, *13*(4), 367-380.

Neuman, Y., & Weizman, E. (2003). The role of text representation in students' ability to identify fallacious arguments. *The Quarterly Journal of Experimental Psychology Section A*, *56*(5), 849-864.

Papert, S. (1980). *Mindstorms--Children, Computers, and Powerful Ideas.* Basic Books, Inc. Publishers, New York.

Paul, R., & Elder, L. (2009). *The miniature guide to critical thinking (5th ed.).* Dillon Beach, CA: The Foundation for Critical Thinking. Retrieved from https://www.criticalthinking.org/files/Concepts_Tools.pdf

Peter, E. E. (2012). Critical thinking: Essence for teaching mathematics and mathematics problem-solving skills. *African Journal of Mathematics and Computer Science Research, 5*(3), 39-43. doi: 10.5897/AJMCSR11.161

Pi, Z., Hong, J., & Yang, J. (2017). Does instructor's image size in video lectures affect learning outcomes?. *Journal of Computer Assisted Learning*, *33*(4), 347-354.

Pokhrel, S., & Chhetri, R. (2021). A literature review on impact of COVID-19 pandemic on teaching and learning. *Higher Education for the Future*, *8*(1), 133-141. Retrieved from https://doi.org/10.1177/2347631120983481

Ramasamy, S. (2011). An Analysis of Informal Reasoning Fallacy and Critical Thinking Dispositions among Malaysian Undergraduates. *Online Submission*. Retrieved from https://files.eric.ed.gov/fulltext/ED525513.pdf

Richardson, J., & Ice, P. (2010). Investigating students' level of critical thinking across instructional strategies in online discussions. *Internet and Higher Education*, 13, 52-59. Retrieved from http://dx.doi.org/10.1016/j.iheduc.2009.10.009

Salina, L., Ruffinengo, C., Garrino, L., Massariello, P., Charrier, L., Martin, B., Favale, M. S., …Dimonte, V. (2012). Effectiveness of an educational video as an instrument to refresh and reinforce the learning of a nursing technique: A randomised controlled trial. *Perspectives on medical education, 1*(2), 67-75.

Schick, T., & Vaughn, L. (2020). *How to think about weird things: Critical thinking for a new age.* New York, NY: McGraw Hill.

Scriven M., & Paul, R. (2007). *Defining critical thinking.* The Critical Thinking Community: Foundation for Critical Thinking. Retrieved from https://www.crc.losrios.edu/files/cassl/DefiningCriticalThinkingc2.pdf

Solon, T. (2007). Generic Critical Thinking Infusion and Course Content Learning in Introductory Psychology. *Journal of Instructional Psychology*, *34*(2).

Stockwell, B. R., Stockwell, M. S., Cennamo, M., & Jiang, E. (2015). Blended learning improves science education. *Cell*, *162*(5), 933-936.

Sundhu, R., & Kittles, M. (2016). Precision teaching: Does training by educational psychologist have an impact? *Educational Psychology in Practice*, *32*(1), 13-23. doi: 10.1080/02667363.2015.1094651

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Measuring cognitive load. In *Cognitive load theory* (pp. 71-85). Springer, New York, NY.

Syed, A. M., Ahmad, S., Alaraifi, A., & Rafi, W. (2020). Identification of operational risks impeding the implementation of eLearning in higher education system. *Education and Information Technologies*, *26*(1), 655-671. doi:10.1007/s10639-020-10281-6

Tiruneh, D. T., De Cock, M., & Elen, J. (2018). Designing learning environments for critical thinking: Examining effective instructional approaches. *International Journal of Science and Mathematics Educucation, 16,* 1065–1089. Retrieved from https://doi.org/10.1007/s10763-017-9829-z

Weinstock, M., Neuman, Y., & Tabak, I. (2004). Missing the point or missing the norms? Epistemological norms as predictors of students' ability to identify fallacious arguments. *Contemporary Educational Psychology*, *29*(1), 77-94.

World Health Organization. (2020, March 20). *WHO Director-General's opening remarks at the media briefing on COVID-19* [Press release]. Retrieved from

https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020

Yousef, A. M. F., Chatti, M. A., & Schroeder, U. (2014). Video-based learning: A critical analysis of the research published in 2003-2013 and future visions. In *eLmL 2014, The Sixth International Conference on Mobile, Hybrid, and On-line Learning* (pp. 112-119).

Author Biography

Dr Angel Tan is a Lecturer in the Department of Psychology at Birmingham City University. She completed her PhD in the Department of Psychology at Edge Hill University. Her PhD focuses on developing a technology-enhanced intervention for higher education learning based on behavioural and pedagogical approaches.

Jean Linda Davies is a Senior Lecturer in the Department of Psychology at Edge Hill University. Her research is in the area of Pedagogical Practice in Higher Education.

Professor Rod Nicolson is a Professor of Psychology at Edge Hill University. His research field is human learning, neuroplasticity and the cerebellum, both for typical and atypical development from young children to older adults.

Dr Themis Karaminis is a Senior Lecturer in the Department of Psychology at Edge Hill University. His research interests are in individual differences in cognitive development, neurodevelopmental disorders, and human learning from developmental, neurocognitive, computational, and educational perspectives.

Compliance with Ethical Standards

Disclosure of Potential Conflicts of Interest

There are no conflicts of interest to report.

Ethics Approval

Informed consent was obtained from all participants in the study and all procedures were approved by the Ethics Committee of the Department of Psychology at Edge Hill University.

Funding Statement

The work was supported by the Department of Psychology at Edge Hill University under the Graduate Teaching Assistantships scheme awarded to the first author.

Availability of Data and Materials

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.