



Addressing Mind Wandering in Video-Based Learning: A Comparative Study on the Impact of Interpolated Testing and Self-explanation

Daniel Ebbert^(✉) , Alrike Claassen , Natasha Wilson , Srecko Joksimovic ,
Negin Mirriahi , and Shane Dawson

University of South Australia, Adelaide, Australia
daniel.ebbert@mymail.unisa.edu.au

Abstract. Task-unrelated thoughts are a common experience for students that negatively impact learning outcomes. In the context of learning from videos, research has attempted to mitigate the impact by including interpolated testing at pauses in the video. The results of these studies are mixed and indicate that interpolated testing may not have a practical effect on reducing task-unrelated thoughts. In the present study, we aim to investigate whether writing self-explanations has a stronger effect than interpolated testing on reducing task-unrelated thoughts and improving learning outcomes. We recruited 138 participants distributed across three groups and presented them with the same video to review. The first group was the control group. The participants in the second group answered interpolated tests, and the participants in the third group wrote self-explanations at pauses in the video. All participants completed a knowledge test before and after watching the video to compare learning outcomes between groups. While previous studies have captured thought reports using probes at pre-determined intervals, we used self-caught thought reports in anticipation that participants expecting to write self-explanations would engage in metacognitive monitoring and thus become meta-aware. We found no significant difference between the groups regarding knowledge gain or meta-awareness. However, the number of thought reports written correlated positively with knowledge gain, indicating that learners who are meta-aware while learning from a video perform better on subsequent knowledge tests.

Keywords: Video-Based Learning · Mind Wandering · Interpolated Testing · Self-Explanation

1 Introduction

Videos are a primary medium for the delivery of course content in higher education. The use of video brings numerous advantages for students in terms of flexible access and timing of study. However, students are also challenged in maintaining focus and minimizing task-unrelated thoughts while engaging with video content. For instance, students may find themselves distracted by thoughts of social engagements or other

non-academic matters. These kinds of thoughts can result in missing information that is presented in the video. While it is inevitable that these off-task thoughts occur, their impact on learning can be mitigated.

Research has shown that while students engage in an educational activity, they will have unrelated thoughts for around 30% of the time [1]. Such task-unrelated thoughts have been shown to negatively correlate with learning outcomes [1, 2]. In the context of learning from videos, practical approaches to reduce the volume of task-unrelated thoughts have been studied to determine impact on learning outcomes. Such methods include the use of notetaking [3, 4], pretesting [5, 6], and interpolated testing [5, 7, 8]. However, the findings from these studies are mixed and inconclusive. For example, interpolated testing, including multiple-choice tests at pauses in the video, has been touted as a means to reduce task-unrelated thoughts. However, the effectiveness of this approach in reducing such thoughts and enhancing learning outcomes remains uncertain with studies producing mixed findings [5]. An alternate method involves prompting students to write self-explanations at specific pauses during a video [9]. This approach shows potential for reducing task-unrelated thoughts, as the writing of self-explanations requires the integration of newly acquired information with existing knowledge [10]. This integration process requires frequent metacognitive monitoring [11], which could lead learners to an increased awareness of having task-unrelated thoughts [12].

In order to test this assumption, we compared the effect of interpolated testing and writing self-explanations to reduce task-unrelated thoughts and increase learning outcomes using self-caught thought reports as a result of meta-awareness.

2 Background

2.1 Conceptualization of Off-Task Thoughts

While engaging in a learning task, students will often experience thoughts unrelated to the goal. These thoughts are called off-task thoughts. We conceptualize off-task thoughts as any internal thoughts unrelated to the task. We can break these off-task thoughts into two sub-concepts: task-unrelated thoughts and task-related interference [13, 14]. For the context of this paper, we conceptualize *task-unrelated thoughts* (TUT) as off-task thoughts that bear no relationship to the intended learning task—for instance, social or non-academic topics. The second type of off-task thought termed *task-related interference* (TRI), is conceptualized as an off-task thought that bears some relationship to the requested learning task. These include thoughts about the appraisal of the specific task or associated criteria [13, 14]. Such as, remembering previous parts of a lecture or wondering how long a lecture video will last.

2.2 Off-Task Thoughts and Learning

Off-task thoughts or task-unrelated thoughts (TUTs) have been shown to interrupt a learners' ability to process external information [15]. These interruptions in information processing detract a student's capacity for engaging with task-related information, thereby decreasing the achievement of learning outcomes [16, 17]. However, the extent

TUTs influence learning outcomes can vary depending on the type of off-task thought a learner engages with. A meta-analysis by Wong et al. [1] on the occurrence of TUTs during educational activities, demonstrated that TUTs occur about 30% of the time and there is a negative relationship between TUT frequency and learning outcomes [1, 2]. While there is no meta-analysis on task-related interference (TRI) during educational activities, TRI is estimated to occur about 20% to 25% of the time and shows no clear correlation with learning outcomes [18]. Due to the negative effect of TUTs on learning outcomes, there have been attempts by various researchers to reduce TUTs, for example, by including interpolated testing at pauses in a video.

2.3 Effect of Interpolated Testing on Task-Unrelated Thoughts

Research on the effect of interpolated testing on TUT when learning from videos has produced mixed results (see Szpunar et al. [7], Jing et al. [8], and Welhaf et al. [5]). An initial study by Szpunar et al. [7] on interpolated testing and TUTs involving a 21-min statistics video found that interpolated testing reduced TUTs, with participants taking more notes than a control group with no tests. The authors reproduced these results of reducing TUTs in a follow-up study [7] with an additional experimental condition in which the participants received the same interpolated test questions and the answers. Similar to the initial study, the participants were allowed to take notes, and the participants in the interpolated testing group took more notes than those in the other two groups [7]. However, subsequent research by Jing et al. [8] using a longer video (40 min) failed to reproduce the effect of reducing TUTs by interpolated testing. In this instance, the interpolated testing group took more notes and performed better on the posttests compared to the other group, which was restudying the material instead of being quizzed. In a follow-up study [8], the participants in the interpolated testing group performed better on the posttests and reported fewer TUTs than the other group. In contrast to the prior studies, the groups did not differ in the quantity of notetaking.

The findings from Jing et al. [8] leads to questions about the relationship between notetaking behavior and TUTs. To address this Welhaf et al. [5] recruited 195 university students to watch a 52-min video in a 2x2 design with the conditions being matched or mismatched pretesting and interpolated testing or restudying. The participants in the interpolated testing condition experienced fewer TUTs than in the restudy condition. However, interpolated testing also did not lead to an increase in learning outcomes. The mixed findings suggest that the desired effect is too small to be of practical use [5]. While the studies by Szpunar et al. [7] and Jing et al. [8] support the use of interpolated testing to reduce TUTs, the participants in these studies were allowed to take notes, and in one study, Jing et al. [8] could not reproduce the effect of reducing TUTs by interpolated testing. When Welhaf et al. [5] prevented participants from taking notes during their study, the effect of interpolated testing on reducing TUTs was diminished but still observable. Considering the varied outcomes associated with interpolated testing, exploring alternative interpolated activities that bolster student learning warrants attention. Investigating whether these activities similarly mitigate TUTs, such as writing self-explanations, could offer valuable insights.

2.4 Self-explanations

A self-explanation is a learning activity in which learners write explanations to themselves. In a study by Fiorella et al. [9], the learning benefits of self-explanations were compared to the effect of rewatching a video or drawing. The findings demonstrated that participants who wrote self-explanations performed better on retention and transfer posttests than the remaining 2 groups. In their follow-up research, Fiorella [19] next compared the effect of writing self-explanations at pauses in a video on learning outcomes to a restudy group. Although differing to the previous study [9] this data collection was performed online, and not in a controlled laboratory environment. The results reflected Fiorella's earlier work [9]. Learners who wrote a self-explanation performed better on a transfer posttest than those who restudied the provided material. There was no difference between the groups on the retention test [19]. These findings strengthen the claim that self-explanations are an effective learning strategy. It remains unclear if including self-explanations during pauses in a video also reduces TUTs.

Given the effectiveness of self-explanation for enhancing understanding and performance in individuals who use self-explanation to address TUTs, it's plausible that anticipating the composition of a self-explanation could also reduce TUTs for participants. The writing of self-explanations requires the learner to engage in metacognitive monitoring [11], which may also be initiated in anticipation of composing the self-explanation while viewing the video. During this metacognitive monitoring, learners may become meta-aware of off-task thoughts.

2.5 Meta-awareness

Meta-awareness is a concept rooted in metacognition, referring to the ability to consciously recognize one's mental state [20, 21]. This process involves periodically turning the mind's focus onto itself. A practice known as metacognitive monitoring or meta-awareness. Essentially, meta-awareness represents reflecting on your individual thoughts, subject to the same limitations inherent in cognition, such as capacity constraints. This limitation implies that an individual may only become aware of certain thoughts while remaining oblivious to others [22]. When one achieves meta-awareness, it often becomes the prevailing thought [12]. However, even partial awareness can suffice to recognize off-task thoughts and redirect attention to the requested learning task.

Researchers usually measure meta-awareness by employing self-caught thought reports [23]. That is, researchers' direct participants to report their off-task thoughts as soon as they become aware. This approach contrasts with the often-used probe-caught method, where participants are interrupted with a thought probe and asked to respond [24]. Researchers have predominantly used the probe-caught method in prior studies examining the effect of interpolated testing on TUTs [5, 7, 8]. It is unknown whether the effect of interpolated testing reduces TUTs when using self-caught thought reports.

2.6 Theoretical Framework

The theoretical framework for this study is grounded in metacognition, focusing on meta-awareness and generative learning theory. Metacognition encompasses the understanding and regulation of one's cognitive processes. An essential aspect of metacognition

is the cycle of metacognitive monitoring and control whereby individuals first monitor their current state of mind and then engage in a control action to guide their cognitive process [25–27].

Generative learning theory complements this understanding. While engaging with a generative learning activity, such as writing a self-explanation, learners need to select relevant information from the content they are studying, organize the selected information into a mental model, and relate this new mental model to prior knowledge, thus integrating the new content with previous knowledge [10]. This process requires learners to engage in metacognitive monitoring and control cycle to assess their current understanding and regulate their cognitive processes to deepen their understanding [11] to positively affect their learning outcomes [10]. In essence, if learners are aware that they are required to write self-explanations during pauses in the instructional video, they may engage in intermittent metacognitive monitoring to ensure a thorough understanding of the content before writing the self-explanation. This practice of intermittent monitoring could support learners to become meta-aware of their off-task thoughts, potentially reducing the frequency of TUTs thereby enhancing more sustained cognitive engagement with the learning material.

While previous studies have documented the occurrence of off-task thoughts during video-based learning and their potential to influence learning outcomes negatively, the results of studies aiming to reduce off-task thoughts using interpolated testing have been mixed. Particularly, the role of anticipatory strategies—such as the foreknowledge of interpolated testing or self-explanation requirements— in promoting meta-awareness and reducing off-task thought occurrences remains unexplored. This gap underscores the need for empirical investigation into how the anticipation of active engagement with learning material influences learners' cognitive processes and, by extension, their learning outcomes. This study aims to examine whether interventions like interpolated testing and self-explanation writing can improve learning outcomes and explore whether anticipating such tasks enhances a learners' meta-awareness.

3 Research Question and Hypotheses

The first research question is: Does writing self-explanations at pauses during video learning sessions more effectively reduce task-unrelated thoughts and enhance learning outcomes compared to interpolated testing or no intervention?

Based on this research question, we formulated the following hypotheses:

Hypothesis 1: There will be a significant difference in knowledge gain among the three experimental groups (control, interpolated testing, self-explanation). Specifically, the self-explanation group will exhibit the highest knowledge gain, followed by the interpolated testing group and the control group.

Hypothesis 2: There will be a significant difference in the number of thought reports written among the three experimental groups (control, interpolated testing, self-explanation). Specifically, the self-explanation group will write the most thought reports, followed by the interpolated testing group, and lastly, the control group.

Given that TUTs are negatively correlated with learning outcomes following the recent meta-analysis by Wong et al. [1] and that TUTs with meta-awareness do not negatively affect subsequent performance [28], the second research question is: Does the number of thought reports, in total and by thought type, correlate with learning outcomes?

4 Method

The University Human Research Ethics Committee granted approval and the study is preregistered at: <https://osf.io/4dg5u>. The data and materials supporting this study are accessible at: <https://data.unisa.edu.au/dap/Dataset.aspx?DatasetID=818303>.

4.1 Design

We employed a between-subjects factorial design with three groups. The three groups correspond to the three experimental conditions. The groups are (1) a control group, (2) an interpolated testing group, and (3) a self-explanation group. Of interest is the intervention used during the video-watching session (answering a thought probe in the control group, answering a thought probe and interpolated tests, or answering a thought probe and writing self-explanations). The control group was also interrupted in their learning to account for uninterrupted study time across groups. An interruption in the study flow can influence the processing of task-related information, which can lead to a change in learning outcome [29, 30]. Uninterrupted study time across groups is controlled by ensuring consistency for all groups. Instead of an activity (interpolated testing or self-explanation), the control group is only asked to answer a thought probe upon interruption. Each participant completed a pre and post-knowledge test as a measure of performance. The study used self-caught thought reports to measure meta-awareness. Self-caught thought reports require participants to indicate their off-task thoughts when they realize they were off-task, which is a form of experience sampling [23, 24].

4.2 Sample Size

The study included 180 participants randomly assigned to one of three experimental groups (control, interpolated testing, and self-explanation). Each group contained 60 participants. Initially, the sample size was based on the study by Szpunar et al. [7], who employed a similar design by comparing TUT rates across three groups. Power analysis with an effect size of $f = .44$ for the difference between groups on TUT frequency [7], $\alpha = .05$, $1 - \beta = .95$, and three groups indicated a required sample size of 28 per group. To leave room for error during data collection, the researchers aimed to recruit 30 participants per group, resulting in 90 participants.

Following the recruitment of the initial 90 participants and subsequent hypotheses testing, which led to their rejection, we decided to deviate from the preregistration protocol and double our sample size to ensure the study was adequately powered. This decision brought our total number of participants to 180.

4.3 Participants

For this study, we recruited 180 participants using the online recruitment platform Prolific, of which 42 did not complete all parts of the study and were excluded. All participants were at least 18 years old, spoke English as their first language, enrolled as a student, and have no diagnosed mental health condition. The participants were required to be students as the study focus was on the use of interpolated testing or self-explanation writing in higher education. The requirement of no mental health condition was included because some mental conditions can lead to an increase in off-task thoughts [31]. Participation in the study was approximately 30 min, and participants received £4.50 as compensation for their time. The participants had an average age of $M = 27$ years ($SD = 8$) with a gender distribution of 34% female and 66% male.

4.4 Procedure

First, participants provided demographic information and completed a pretest of 5 knowledge questions. Following the pretest, off-task thought was operationalized to the participants as any thoughts not focused on the current task and the video content at that moment. A short comprehension quiz assessed participants' understanding of off-task thought. Participants were instructed to report their off-task thoughts whenever they realized they were off task by clicking on a button that opened a pop-up where the participants could record what they were thinking.

After this explanation, the participants were informed about the upcoming learning activity (interpolated testing or self-explanations) at pauses in the video. All participants then watched the same video titled "Why Is Blue So Rare In Nature?" [32] with a duration of 8 min and 20 s. This video explores why blue is rare in living organisms despite its prevalence in elements like the sky and ocean. The participants could not control the video; pausing or navigating within the video player was disabled.

The researchers interrupted the video watching of all three groups three times, with each interruption occurring at the same time within the video for all groups. The first interruption occurred at 2 min and 33 s, the second at 5 min and 2 s, and the third at 7 min and 30 s. The main difference between the groups was what the researchers asked the participants to do during each interruption. The control group was requested to write a thought report, the interpolated testing group to answer three multiple-choice questions, and the self-explanation group to write self-explanations. After watching the video, the researchers instructed the participants to engage in a filler task to reduce information retrieval from working memory. This step involved completing a math test in a manner similar to that adopted by Szpunar et al. [7] and Jing et al. [8]. Upon completing the filler task, the researchers asked the participants to answer the posttest, which consisted of the same knowledge test as the pretest. The participants had not received the results of the pretest and thus had to obtain the information required to answer the posttest from the video.

4.5 Data Coding

The first two authors of this paper manually coded the free-text thought reports provided by the participants. The coding categories were on-task, external distraction, TUT, and

TRI, following the conceptual division of cognitive experiences by Smallwood and Schooler [33]. A participant's thought about the video's content at the time of the thought report, was classified as on-task. In contrast, a thought report about external perceptions (e.g., pets in the room) was deemed an external distraction. Thoughts unrelated to the task (watching the video) were classified as TUT. We classified a thought about the task that represented an interruption in the flow of the task (e.g., thinking about the task's difficulty, questioning understanding) as TRI. All thought reports were coded independently by two coders. The interrater agreement between the two coders, measured using Cohen's kappa, was $\kappa = 0.47$. The coders resolved disagreements through discussion.

4.6 Analysis Plan

Following the preregistration protocol, we planned to test hypothesis 1 using an Analysis of Covariance, and hypothesis 2 using an Analysis of Variance. We also planned to conduct an exploratory analysis using a Multivariate Analysis of Variance to determine group differences by thought type. Additionally, we planned to use a Kruskal-Wallis test instead if the assumptions for either of the parametric analysis methods were not met. Additionally, in deviation from the preregistration protocol, we performed a correlational analysis between thought type frequency and learning outcomes to answer the second research question.

5 Results

5.1 Data Exclusion

Incomplete data was excluded from the analysis. Of the 180 participants recruited, 42 did not complete all parts of the study and were excluded. This resulted in 138 participants for analysis. Before each analysis segment, an outlier detection process was carried out. Outliers were identified using the standard deviation method, a cutoff of two standard deviations, and the interquartile range method. All outliers were then excluded from the analysis. The resulting group distribution is 47 participants in the control group, 51 in the interpolated testing group, and 40 in the self-explanation group.

5.2 Hypothesis 1 Testing

A Kruskal-Wallis test was conducted to assess the groups' differences in knowledge gained on the posttest. The Kruskal-Wallis test was chosen because the assumptions for an Analysis of Covariance were unmet. Table 1 shows the group's pretest, posttest, and change score (posttest result minus pretest result). The results of the Kruskal-Wallis test did not indicate a statistically significant difference between the groups, $H(2) = 1.913$, $p = .384$. Therefore, Hypothesis 1 is rejected.

Table 1. Knowledge test scores by group

Variable	Control (n = 47) Mean (SD)	Interpolated Testing (n = 51) Mean (SD)	Self-explanation (n = 40) Mean (SD)
Pretest	0.85 (0.81)	0.76 (0.86)	0.93 (0.86)
Posttest	2.87 (1.19)	3.12 (1.01)	3.33 (1.25)
Pre/Posttest change	2.02 (1.41)	2.35 (1.29)	2.40 (1.48)

5.3 Hypothesis 2 Testing

A Kruskal-Wallis test was conducted to test whether there was a difference between the groups in the number of self-caught thought reports they wrote. The Kruskal-Wallis test was chosen because the assumptions for an Analysis of Variance were unmet. The results of the Kruskal-Wallis test did not indicate a statistically significant difference between the groups, $H(2) = 3.453$, $p = .178$. Therefore, Hypothesis 2 is rejected.

5.4 Exploratory Analysis

In addition to the previously described confirmatory analysis, an exploratory analysis was undertaken to determine differences between the groups and the number of thought reports written for each type of off-task thought (TUT or TRI). For this analysis, the self-caught thought reports and the probe-caught thought reports were used. An overview of the number of thought reports by type is presented in Table 2.

Table 2. Number of thought reports by type and group

Variable	Control (n = 294)	Interpolated Testing (n = 321)	Self-explanation (n = 260)
On Task	38	57	39
External Distraction	15	34	14
TUT	127	119	115
TRI	114	111	92

The assumptions for a Multivariate Analysis of Variance to compare the number of thought reports by type and group were not met. Instead, the Kruskal-Wallis test was used to determine whether groups had statistically significant differences for each type of off-task thought. The Kruskal-Wallis tests revealed no statistically significant differences among groups for TUT, $H(2) = 0.546$, $p = .761$, as well as for TRI, $H(2) = 2.522$, $p = .283$.

5.5 Correlational Analysis

To answer the second research question, we conducted a correlational analysis between the total number of thought reports, TUT, TRI, and learning outcomes (posttest minus pretest result). The assumptions for Pearson's correlation were not met; instead, Kendall's rank correlation was used. The number of thought reports a participant wrote positively correlates with the learning outcome, $\tau = .1527$, $H(1) = 2.33$, $p = .009892$. The number of TUTs showed no correlation with learning outcome, $\tau = .07999$, $H(1) = 1.198$, $p = .8845$, and neither did the number of TRIs, $\tau = .06661$, $H(1) = 0.995$, $p = .3197$.

6 Discussion

This study compared the effects of interpolated testing and writing self-explanations at pauses in a video on learning outcomes and the number of thought reports written by participants. The first hypothesis stated that there would be a significant difference between the three groups (control, interpolated testing, and self-explanation) in the result of a knowledge test. This hypothesis was rejected based on the result of a Kruskal-Wallis test, showing no significant difference between the groups. The findings indicate that the interpolated testing group did not perform better than the control group, which is in line with prior research. Welhaf et al. [5] also found no increase in learning outcomes for participants who were prompted with interpolated testing while learning from a video. However, the results of the present study and that by Welhaf et al. [5] lies in contrast to the reported research by Szpunar et al. [7] and Jing et al. [8] who both allowed their participants to take notes and found significant effects for increased learning outcomes due to interpolated testing. Notetaking was not allowed in the study by Welhaf et al. [5] and in our study and in both cases no significant effect of interpolated testing on increased learning outcomes could be observed.

Considering that our data collection was performed online using a retention test, our results align with previous research. While the first study by Fiorella et al. [9], conducted in a controlled laboratory environment, found a significant effect for writing self-explanations on a retention and a transfer test [9], the effect of writing self-explanation on a retention test was not found when the data was collected in a less controlled online environment [19]. These results align with the result of our hypothesis test, in both cases the data was collected in an uncontrolled online environment and found no effect of writing self-explanations on a retention test.

The second hypothesis was a direct test of the theoretical framework and stated that there would be a significant difference between the groups in the number of self-caught thought reports they wrote. These self-caught thought reports indicate that a learner engaged in metacognitive monitoring and realized they were off-task, they were meta-aware. The second hypothesis was motivated by the assumption that learners would engage in metacognitive monitoring in anticipation of learning tasks at pauses in a video. In short, learners would become meta-aware and write a thought report if they were off task. This hypothesis was rejected based on a Kruskal-Wallis test between the three groups on the number of thought reports they wrote. Combined with generative

learning theory [11], this result suggests that while learners need to engage in metacognitive monitoring while writing the self-explanation, they do not do so in expectation of having to write a self-explanation or answer a multiple-choice test. This result was further explored in an exploratory analysis by breaking down the thought reports into the categories of TUTs and TRIs and comparing the groups on these types of thoughts. The results of this comparison showed that there also were no differences between the groups in the number of TUTs or TRIs reported. This result shows that the effect of reducing TUT by including interpolated testing in a video can only be observed when measuring TUTs using thought probes [5, 7, 8] and further supports the assertion by Welhaf et al. [5] that using interpolated testing to reduce TUTs might be of limited utility.

Additionally, we explored whether the total number of thought reports, TUTs, or TRIs correlated with learning outcomes. Based on the meta-analysis by Wong et al. [1], the assumption was that TUTs would correlate negatively with learning outcomes. Although the results of the meta-analysis by Wong et al. [1] are mainly based on studies using probes to measure TUT, and in the subgroup analysis for their meta-analysis, there was no difference in the correlation with learning outcomes between probe-caught TUTs and self-caught TUTs, this meta-analysis only included two papers in which the participants learned from videos and used self-caught thought report. The first paper by Kopp et al. [34] resulted in a nonsignificant correlation between off-task thoughts and comprehension performance of $r = -0.024$. The second paper by Zhang et al. [35] included two studies using self-caught thought reports while learning from a video. The first study showed a significant correlation between TUTs and a posttest score of $r = 0.37$, and the second a nonsignificant correlation of $r = 0.04$. Therefore, while the overall result of the meta-analysis shows a negative correlation between TUTs and learning outcomes, the comparable sources paint a different picture with mixed results. Our result of no correlation between TUTs and learning outcomes ($\tau = .07999$) is consistent with prior studies when considering that we used self-caught thought reports to measure off-task thoughts while learning from a video. While there also was no correlation between TRIs and learning outcomes, this was an expected result for TRIs based on the estimation by Welhaf et al. [18]. The only significant finding was that the total number of thought reports correlated positively with increased performance on the knowledge test. This result of a positive correlation of meta-awareness with test performance lies in contradiction to the previous research. Smallwood et al. [28] found that TUTs with meta-awareness did not negatively affect performance [28] and that there is no correlation between meta-awareness and performance [36]. This result means that participants who wrote more thought reports and thus were more meta-aware performed better on the subsequent knowledge test. This indicates that meta-awareness can mitigate the negative consequences of off-task thoughts on learning outcomes.

7 Limitations

A limitation of the current study involves the online data collection methodology employed via Prolific. This approach prevented the researchers from regulating the environment in which participants engaged with the study. Despite instructions for participants to undertake the study in a distraction-free setting, we cannot exclude the potential for engagement in environments prone to interruptions. Such conditions may have

impacted the number of thought reports documenting external distractions, accounting for 11% of all thought reports. The online nature of the study further limited control over whether participants took notes during the experiment. Notetaking reduces the occurrence of TUTs in the expectations of interpolated testing [7, 8]. Consequently, it is possible that some participants ignored instructions and took notes while watching the videos, leading to increased performance on the posttest. This limitation could be overcome by replicating this study design in a controlled environment.

The duration of the video utilized in this research (8 min and 20 s) is notably shorter than those used in prior studies investigating the impact of interpolated testing on TUTs, which have spanned from 21 to 52 min [5, 7, 8]. This discrepancy raises the possibility that the abbreviated video length contributed to the absence of observable differences in the volume of thought reports between groups. Considering that TUTs increase over time [37], the chosen video might have been too short for many off-task thoughts to occur, which could be overcome by using a longer video. Moreover, the video's format may have served as an additional variable. Unlike previous research that employed lecture videos from higher education contexts, this study utilized an engaging educational YouTube video, which could have contributed to the participants' paying closer attention to the content and could be a topic for future research.

Lastly, the employment of Open edX's experimental function, in conjunction with pre-established accounts on a research and learning platform distributed to Prolific users, coupled with a high attrition rate, resulted in unequal group distributions. This imbalance may have further influenced study findings and could be overcome by recruiting participants by group until the required number of participants by group is reached.

8 Conclusion

This study investigated the impact of interpolated testing and writing self-explanations at video pauses on learning outcomes and participant engagement in metacognitive monitoring, as evidenced by thought reports. Contrary to the stated hypotheses, we found no significant differences in knowledge test results across the control, interpolated testing, and self-explanation groups. This outcome challenges previous findings on the efficacy of these strategies for enhancing learning from video content. Moreover, our investigation into the frequency of self-caught thought reports did not reveal the anticipated differences among groups, suggesting that these learning strategies may not influence learners' metacognitive monitoring to the extent hypothesized. Further exploration into the types of thought reports—TUTs and TRIs—also showed no significant differences among the groups, raising questions about the utility of interpolated testing in reducing off-task thoughts.

Interestingly, a positive correlation between the total number of thought reports and improved knowledge test performance suggests that higher meta-awareness levels could counteract the potential learning detriments of off-task thoughts. While it is inevitable that learners will experience off-task thoughts, whether they are aware of their off-task thoughts or not could be the factor that influences if these off-task thoughts also influence their learning outcomes. Consequently, this calls for further research into how meta-awareness while learning can be increased.

Acknowledgments. The first author is supported by an Australian Government Research Training Program international (RTPi) Scholarship.

Disclosure of Interests. The authors have no competing interests to declare relevant to this article's content.

References

1. Wong, A.Y., Smith, S.L., McGrath, C.A., Flynn, L.E., Mills, C.: Task-unrelated thought during educational activities: A meta-analysis of its occurrence and relationship with learning. *Contemp. Educ. Psychol.* **71**, 102098 (2022). <https://doi.org/10.1016/j.cedpsych.2022.102098>
2. Bonifacci, P., Viroli, C., Vassura, C., Colombini, E., Desideri, L.: The relationship between mind wandering and reading comprehension: A meta-analysis. *Psychon. Bull. Rev.* **30**, 40–59 (2023). <https://doi.org/10.3758/s13423-022-02141-w>
3. Kane, M.J., Smeekens, B.A., von Bastian, C.C., Lurquin, J.H., Carruth, N.P., Miyake, A.: A combined experimental and individual-differences investigation into mind wandering during a video lecture. *J. Exp. Psychol. Gen.* **146**, 1649–1674 (2017). <https://doi.org/10.1037/xge0000362>
4. Wong, S.S.H., Lim, S.W.H.: Take notes, not photos: Mind-wandering mediates the impact of note-taking strategies on video-recorded lecture learning performance. *J. Exp. Psychol. Appl.* **29**, 124–135 (2023). <https://doi.org/10.1037/xap0000375>
5. Welhaf, M.S., Phillips, N.E., Smeekens, B.A., Miyake, A., Kane, M.J.: Interpolated testing and content pretesting as interventions to reduce task-unrelated thoughts during a video lecture. *Cogn. Research.* **7**, 26 (2022). <https://doi.org/10.1186/s41235-022-00372-y>
6. Pan, S.C., Schmitt, A.G., Bjork, E.L., Sana, F.: Pretesting reduces mind wandering and enhances learning during online lectures. *J. Appl. Res. Mem. Cogn.* **9**, 542–554 (2020). <https://doi.org/10.1016/j.jarmac.2020.07.004>
7. Szpunar, K.K., Khan, N.Y., Schacter, D.L.: Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6313–6317 (2013). <https://doi.org/10.1073/pnas.1221764110>
8. Jing, H.G., Szpunar, K.K., Schacter, D.L.: Interpolated testing influences focused attention and improves integration of information during a video-recorded lecture. *J. Exp. Psychol. Appl.* **22**, 305–318 (2016). <https://doi.org/10.1037/xap0000087>
9. Fiorella, L., Stull, A.T., Kuhlmann, S., Mayer, R.E.: Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *J. Educ. Psychol.* **112**, 895–906 (2020). <https://doi.org/10.1037/edu0000408>
10. Fiorella, L., Mayer, R.E.: Learning by Self-Explaining. In: *Learning as a Generative Activity: Eight Learning Strategies that Promote Understanding*. pp. 124–150. Cambridge University Press (2015). <https://doi.org/10.1017/CBO9781107707085>
11. Fiorella, L., Mayer, R.E.: Introduction to Learning as a Generative Activity. In: *Learning as a Generative Activity: Eight Learning Strategies that Promote Understanding*. pp. 1–19. Cambridge University Press (2015). <https://doi.org/10.1017/CBO9781107707085>
12. Chin, J.M., Schooler, J.W.: Meta-Awareness. In: *Encyclopedia of Consciousness*. pp. 33–41. Elsevier (2009). <https://doi.org/10.1016/B978-012373873-8.00051-7>
13. Smallwood, J., Baracaia, S.F., Lowe, M., Obonsawin, M.: Task unrelated thought whilst encoding information. *Conscious. Cogn.* **12**, 452–484 (2003). [https://doi.org/10.1016/S1053-8100\(03\)00018-7](https://doi.org/10.1016/S1053-8100(03)00018-7)

14. Smallwood, J., Obonsawin, M., Heim, D.: Task unrelated thought: The role of distributed processing. *Conscious. Cogn.* **12**, 169–189 (2003). [https://doi.org/10.1016/S1053-8100\(02\)00003-X](https://doi.org/10.1016/S1053-8100(02)00003-X)
15. Kam, J.W.Y., et al.: Slow Fluctuations in Attentional Control of Sensory Cortex. *J. Cogn. Neurosci.* **23**, 460–470 (2011). <https://doi.org/10.1162/jocn.2010.21443>
16. Smallwood, J., Fishman, D.J., Schooler, J.W.: Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychon. Bull. Rev.* **14**, 230–236 (2007). <https://doi.org/10.3758/BF03194057>
17. Smallwood, J.: Mind-wandering While Reading: Attentional Decoupling, Mindless Reading and the Cascade Model of Inattention. *Language and Linguistics Compass.* **5**, 63–77 (2011). <https://doi.org/10.1111/j.1749-818X.2010.00263.x>
18. Welhaf, M.S., et al.: An exploratory analysis of individual differences in mind wandering content and consistency. *Psychol. Conscious. Theory Res. Pract.* **7**, 103–125 (2020). <https://doi.org/10.1037/cns0000180>
19. Fiorella, L.: Learning by explaining after pauses in video lectures: Are provided visuals a scaffold or a crutch? *Appl. Cogn. Psychol.* **36**, 1142–1149 (2022). <https://doi.org/10.1002/acp.3994>
20. Schooler, J.W., Smallwood, J., Christoff, K., Handy, T.C., Reichle, E.D., Sayette, M.A.: Meta-awareness, perceptual decoupling and the wandering mind. *Trends Cogn. Sci.* **15**, 319–326 (2011). <https://doi.org/10.1016/j.tics.2011.05.006>
21. McComas, W.F.: Metacognition. In: McComas, W.F. (ed.) *The Language of Science Education*, pp. 63–63. SensePublishers, Rotterdam (2014). https://doi.org/10.1007/978-94-6209-497-0_55
22. Schooler, J.W.: Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends Cogn. Sci.* **6**, 339–344 (2002). [https://doi.org/10.1016/S1364-6613\(02\)01949-6](https://doi.org/10.1016/S1364-6613(02)01949-6)
23. Chu, M.T., Marks, E., Smith, C.L., Chadwick, P.: Self-caught methodologies for measuring mind wandering with meta-awareness: A systematic review. *Conscious. Cogn.* **108**, 103463 (2023). <https://doi.org/10.1016/j.concog.2022.103463>
24. Weinstein, Y.: Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behav Res.* **50**, 642–661 (2018). <https://doi.org/10.3758/s13428-017-0891-9>
25. Nelson, T.O.: Metamemory: A Theoretical Framework and New Findings. *Psychol. Learn. Motiv.* **26**, 125–173 (1990). [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
26. Nelson, T.O., Narens, L.: Why Investigate Metacognition? In: Metcalfe, J. and Shimamura, A.P. (eds.) *Metacognition: Knowing about Knowing*, pp. 1–25. The MIT Press (1994). <https://doi.org/10.7551/mitpress/4561.003.0003>
27. Nelson, T.O.: Consciousness and Metacognition. *Am. Psychol.* **51**, 102–116 (1996). <https://doi.org/10.1037/0003-066X.51.2.102>
28. Smallwood, J., McSpadden, M., Schooler, J.W.: The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychon. Bull. Rev.* **14**, 527–533 (2007). <https://doi.org/10.3758/BF03194102>
29. Wirzberger, M., Esmaili Bijarsari, S., Rey, G.D.: Embedded interruptions and task complexity influence schema-related cognitive load progression in an abstract learning task. *Acta Physiol. (Oxf)* **179**, 30–41 (2017). <https://doi.org/10.1016/j.actpsy.2017.07.001>
30. Conard, M.A., Marsh, R.F.: Interest level improves learning but does not moderate the effects of interruptions: An experiment using simultaneous multitasking. *Learn. Individ. Differ.* **30**, 112–117 (2014). <https://doi.org/10.1016/j.lindif.2013.11.004>
31. Seli, P., Smallwood, J., Cheyne, J.A., Smilek, D.: On the relation of mind wandering and ADHD symptomatology. *Psychon. Bull. Rev.* **22**, 629–636 (2015). <https://doi.org/10.3758/s13423-014-0793-0>

32. Hanson, J.: Why Is Blue So Rare In Nature?, <https://www.youtube.com/watch?v=3g246c6Bv58>
33. Smallwood, J., Schooler, J.W.: The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness. *Annu. Rev. Psychol.* **66**, 487–518 (2015). <https://doi.org/10.1146/annurev-psych-010814-015331>
34. Kopp, K., Mills, C., D’Mello, S.: Mind wandering during film comprehension: The role of prior knowledge and situational interest. *Psychon. Bull. Rev.* **23**, 842–848 (2016). <https://doi.org/10.3758/s13423-015-0936-y>
35. Zhang, H., Miller, K.F., Sun, X., Cortina, K.S.: Wandering eyes: Eye movements during mind wandering in video lectures. *Appl Cognit Psychol.* **34**, 449–464 (2020). <https://doi.org/10.1002/acp.3632>
36. Zedelius, C.M., Broadway, J.M., Schooler, J.W.: Motivating meta-awareness of mind wandering: A way to catch the mind in flight? *Conscious. Cogn.* **36**, 44–53 (2015). <https://doi.org/10.1016/j.concog.2015.05.016>
37. Risko, E.F., Buchanan, D., Medimorec, S., Kingstone, A.: Everyday attention: Mind wandering and computer use during lectures. *Comput. Educ.* **68**, 275–283 (2013). <https://doi.org/10.1016/j.compedu.2013.05.001>