The "How" Matters: Evaluating different Video Types for Cybersecurity MOOCs *

Daniel Köhler, Wenzel Pünter, and Christoph Meinel

Hasso-Plattner-Institute, University Potsdam
Potsdam, Germany
daniel.koehler@hpi.de
wenzel.puenter@hpi.de

Abstract. Teachers and educators are usually required to transfer knowledge to groups of learners simultaneously. However, not all students necessarily learn in the same way. In cybersecurity education, severe differences between understanding and applying knowledge are observed. In our study, we performed Randomized Controlled Trials with more than 1,500 participants to compare different educational videos: a presentation with slides, an interview, and a short animation. We evaluate learning success for the three dimensions of cybersecurity: Perception, Protection, and Behavior and observe that traditional presentations with slides perform best for achieving fundamental understanding (Perception), tested in recall exercises. Animation videos achieve the best learning success in transfer tasks, such as for assessing protective measures. While statistically insignificant, we observe a slight tendency of animation video learners to apply the learned behavior best, while learners of the interview videos performed worst.

Keywords: Video Styles · Online Education · Cybersecurity Awareness · Field Study

1 Introduction

Massive Open Online Courses (MOOCs) have found application worldwide in the last ten years. Enabling teachers and learners to come together, online education technologies have seen a surge in usage during the Covid19-Pandemic.

* — **Authors Version** —
— *Do Not Distribute* —
This preprint has not undergone post-submission improvements or corrections. Accepted and to be published in EC-TEL2023 proceedings by Springer.
The Version of Record is to be found at:
DOI: https://doi.org/10.1007/978-3-031-42682-7_11

O. Viberg et al. (eds.), *Responsive and Sustainable Educational Futures*, Lecture Notes in Computer Science 14200

The availability of MOOCs allows every learner to find a course on almost any topic. With the ongoing digitization of the world, internet-connected devices are used in evergrowing areas of our lives. Unfortunately, all these connected devices pose a range of attack surfaces to cyber criminals. With the increasing penetration of day-to-day computer use, laypersons and smaller companies are also becoming victims of cybercrime [8,3]. Cybersecurity awareness can hence be considered one of the most essential skills to develop in the upcoming years. Online education programs often target knowledge recall with videos and multiple choice quizzes [24]. In the field of cybersecurity awareness, however, different types of skills are required. Awareness requires an inherent sensitization for a topic in addition to the technical understanding and knowledge about protective measures. MOOCs are often mainly based on traditional video formats such as the *Talking Head with Slides* [27]. However, past research has already shown (c.f. Section 2.1) that educators can generate some impact on learner's perceptions by providing them with different types of educational videos[7,16,31,28]. We, therefore, pose the following research question:

Research Question: Which impact on learning outcomes can teachers provide by varying the video type employed in online cybersecurity education?

We provide an overview of past work on different video styles in Section 2.1, followed by an introduction to cybersecurity awareness in Section 2.2. In our real-world field study, we derive honest results by sending the unknowing users phishing emails to assess them for their behavior towards the threat of phishing. Our experiment employs randomization to assign learners to different video types, described in Section 3. Section 4 presents the results of our study and the employed tests for the different dimensions of cybersecurity. We discuss our findings, methodology, and threats to validity in Section 5. We thereby contribute to the body of research on different video types in online education:

Contribution: We provide a real-world study with more than 1,500 participants covering the impact of different video types on learning success for different task types. We identify that the video type does play a significant role in learning success for *Recall* and *Transfer* tasks (c.f. Sections 4.2 and 4.3). Noticeable, but statistically insignificant is the observation that learners who watched *Animation* videos were best in appropriately reacting to phishing emails (c.f. Sec. 4.4).E.g., approximately 3% fewer participants who watched the *Animation* submitted personal information to malicious websites in our field study (*Animation*: 7.4%, *Interview*: 9.3%, *Slides*: 10.7%).

2 Background and Related Work

In most studies, the impact of different video styles is evaluated by success in recall tasks. However, learning can take different levels, e.g., recall and understanding or applying the knowledge in transfer tasks. Various models have been developed to conceptualize the levels of learning, such as, e.g., Bloom's Taxonomy [2], Anderson and Krathwohl's model [20], or Metzger's framework [23].

Our education context of cybersecurity superimposes even different challenges. Concretely, learners must constantly be aware of threats and dangers, even when simply checking their emails. Such unconscious awareness might require additional training compared to other areas of knowledge. In Section 2.2, we present a model to evaluate dimensions of security awareness. Further, we briefly introduce the threat of phishing in Section 2.3.

2.1 Related Work on Video-Styles

In their 2016 literature review, *Santos-Espino et al.* qualitatively analyzed the used video styles of 115 MOOCs [27]. The authors build on several pieces of literature trying to define available video styles, e.g., [10] and [21]. *Santos-Espino et al.* derived seven video styles, such as *Talking Head*, *Live Lectures*, or *Slides*. They further quantitatively analyzed the usage of different styles throughout the MOOCs. The authors report five style combinations, of which *Talking Head with Slides* was the most used in the analyzed courses. In 2014, *Guo et al.* analyzed 6.9 million video-watching sessions, deriving the impact of different video styles on learner success [11]. The authors identified that e.g., shorter videos are much more engaging to learners and that videos of the combination *Talking Head with Slides* are more engaging than *Slides* alone.

Besides categorizing video types, some researchers have already studied the impact of different video types on the learning success of different task types, such as *Recall* and *Transfer*. Studies from as early as 2014 [7,16] or 2017 [31], however were largely inconclusive. While *Kizilcec et al.* [16] observe no difference for recall tasks and motivate to research transfer tasks, *Wang and Antonenko* [31] observe differences only for easy recall tasks, but explicitly not for transfer tasks.

Building on the previous work, in 2022, *Steinbeck et al.* evaluated a field experiment conducted in a German-speaking MOOC with approximately 3000 participants on the impact of video styles on learner perception and learning success [28]. The authors prepared modern *Explainer* videos as seen on platforms such as YouTube to compare against traditional *Talking Head with Slides* videos. The authors observed a different perception of the speaker's focus as well as better scores on *Recall* posttests in their study. Still, they did not identify an impact on scores in *Transfer* tasks.

2.2 Dimensions of Security Awareness

To adequately cover the context of cybersecurity in education technologies, particularly cybersecurity awareness, we build on a definition for security awareness by *Jaeger* [13]: *Security Awareness is a state of mind, derived by education and experience in which persons are capable of understanding and protecting themselves against security threats.* Such cybersecurity awareness is not only related to companies and professional activities of people anymore. Instead, it is similarly required in one's private life.

In 2014, *Hänsch and Benenson* performed a literature review on the scope and dimensions of *Security Awareness* [12]. The authors analyzed more than 25

publications from previous years. They derived three major dimensions towards security awareness: *Perception* covers the awareness and knowledge required to (theoretically) understand that dangers in cyberspace exist. *Protection* covers knowledge on possible security measures users can employ to enhance their security hygiene and increase their protection against cybersecurity threats. However, *Protection* only covers the knowledge about those security measures; e.g., most people know they should use complex passwords. However, the least people do it [9,25]. Hence, the dimension of *Behavior* assesses whether people are applying their knowledge to a reasonable behavior (change). Conceptionally, the dimensions of security awareness can be compared to levels of learning from educational models such as Bloom's Taxonomy [2].

2.3 Phishing Attacks

Phishing describes the malicious practice of tricking individuals into illicit actions such as revealing (confidential) information or performing unintended actions [32]. In security incidents like data breaches, phishing is one of the most prominent vectors for initial access [6].

Usually, attackers send, e.g., a fake email to their targets, presenting an issue that the recipient must react to. One typical example is a supposed full email postbox. Recipients would be asked to click on a link and log in to their postbox to remediate the issue. Prevalent research in phishing susceptibility and cybersecurity awareness shows that while technical measures to increase security exist, there is no absolute protection [14,15,19]. Hence, users — laypersons and professionals alike — need to be educated about the possible dangers. In professional contexts, employees are often offered dedicated training programs. In private contexts, one form of training can be participation in an online course such as our MOOC on cybersecurity. Generally, embedded training programs in which employees are sent fake phishing emails and educated after falling for the bait are reported as most impactful towards increased sensitization among trainees [15].

3 Methodology & Experiment Setting

Our study evaluates different levels of learning success for cybersecurity awareness generated by three different video types presented to learners. We embedded the experiment into a German-speaking Massive Open Online Course on cybersecurity aimed at beginners. The course spanned six weeks (Oct. - Dec. 2022), with the study being performed during week 4 of the course, alongside the respective thematic context of threats from the internet. The first stage of our study features a Randomized Controlled Trial [29], covered in more depth in Section 3.1 for the three types of learning content and the cybersecurity awareness dimensions of *Perception* and *Protection*. In the second phase, all learners who consented to participate in a phishing study were sent three iterations of

phishing emails described in Section 3.3. The phishing study was approved by the Institutional Review Board of the University of Potsdam.

Fundamentally, our study uses a *True Experimental Design* according to principles outlined by Campbell and Stanley [4]. The experiment features both randomization to distribute participants equally and a pre-test to ensure that the distribution is not accidentally biased by other influences.

3.1 Experiment Setup

The fourth week of our online cybersecurity course covered threats from the internet. Twelve videos present various threats and protective measures, respectively.

$$O_P \quad R \quad \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \quad R^* \quad \begin{pmatrix} O_1 \\ O_2 \end{pmatrix} \quad O_M$$

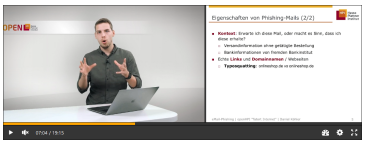
The semi-formal description of our study shows that all learners take a pre-test for knowledge (O_P). After that, they are randomly (round-robin) assigned (R) to either of the three treatments in the form of three different types of learning videos, covered in more detail in Section 3.2:

X_1 : Traditional video with one presenter and slides; *(Talking Head with) Slides*
X_2 : Interview with two presenters, without slides; *Interview*
X_3 : Short animation video; *Animation*

Some users skip the videos and directly go to the following self-tests. We do not take that group of users into account for our analysis of differences between the three treatment groups. To evaluate learning success towards the levels of *Perception* as well as *Protection* of cybersecurity awareness (c.f. Sec. 2.2), we conducted two different self-tests (O_1, O_2) after learners consumed their respective video. Again, learners were randomly (round-robin) assigned (R^*)to either test group. One self-test covered the theoretical understanding of phishing threats (O_1) as a recall exercise. The other covered assessing exemplary emails to identify possible phishing emails (O_2) as a transfer exercise. Finally, all learners who participated in the phishing study (c.f. Section 3.3) received multiple phishing emails over three months (O_M).

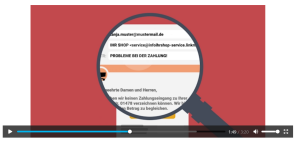
3.2 Video Types

We presented the randomized groups of learners with one of three video types. Figure 1 shows screenshots from all three video types. The treatment X_1, *(Talking Head with) Slides*, is our control group for the default video type employed in most of our online courses. To perform a reliable study, we aimed to ensure that all videos cover the same content. As we included an external animation video, which's content we could not define, we used that as a base for developing



X_1: *Slides*: Traditional video with a presenter and slides

X_2: *Interview:* Two speakers, no slides



X_3: *Animation*: Short, animated video[1]

Fig. 1: Screenshots of the three video types used in the MOOC

the content for the *Slides* and *Interview* videos. Obviously, with the formats of a non-scripted *Slide* presentation and an *Interview*, in both variants, the speakers took longer to present the content. The traditional video with slides and the interview with two presenters were approximately 14 minutes long, and the animation video was 3.5 minutes.

3.3 Phishing Study for Security Awareness

As previously highlighted, the learning goals of security awareness cover three dimensions: *Perception*, *Protection* and *Behavior* (c.f. Sec. 2.2). Phishing studies in which (unknowing) participants are sent phishing emails are one of the most prominent assessment measures for *Behavior* in security awareness [17,15]. During our study, participants were sent multiple phishing emails throughout three months [18]. Each month, we sent one email to participants. To account for missing difficulty between the three emails, we evaluate accumulated responses throughout the emails. Learners would have been able to identify all of the phishing emails because of their missing or wrong context. Notably, the missing context has been one of the main criteria highlighted throughout all three learning videos to identify phishing emails.

3.4 Statistical Analysis

During our analysis, we employ *Kruskal-Wallis* as well as *chi-squared* tests to verify the statistical validity of our results. The Kruskal-Wallis test [22] is a statistical method to analyze non-parametric datasets. It indicates whether the compared samples could originate from the same distribution. A significance in test results confirms that the samples originate from different distributions. The test is applied to categorical input and quantitative output variables. For post hoc analysis, we employ *Dunn Bonferroni Tests*. On the other hand, we use Pearson's chi-squared test [26,5] to compare categorical output variables. By performing chi-squared (χ^2) tests of independence, one verifies that occurrences of variables in different groups are independent of one another. Throughout our

[1] For the German-speaking participants of the MOOC, we embedded an animation video provided by the *Federal Office for Information Security*, available at: https://multimedia.gsb.bund.de/BSI/Video/Sicher_im_Internet/Phishing.mp4

analysis, we employ $\alpha = 0.05$ as a quasi-standard in significance tests. We use Cohen's d [30] to calculate standardized effect sizes.

4 Study Results

4,490 participants engaged in our MOOC during the course start, from which 1,867 were still active in course week four, which contained the content on phishing and our study. Table 1 presents the amounts of participants in each of the respective sections of our study. Our study employed randomization between X_1, X_2, X_3 and between the self-tests O_1, O_2. Learners who chose not to watch any of the education videos are neither listed nor considered for further analysis. Participants who provided consent to receive emails were enrolled into O_M.

Table 1: Number of learners participating in the study, divided for the different types of learning content and the form of learning assessment.

		Participants			
	Learning Content	O_P	O_1	O_2	O_M
X_1	Slides	513	217	198	172
X_2	Interview	517	204	225	189
X_3	Animation	472	219	213	179
	Total	1,502	1,276		540

The following sections discuss the different stages of the experiment in detail. Section 4.1 presents our analysis of the pretest O_P employed to compare the three groups of participants before any treatment. Section 4.2 presents the results for the level of *Perception* (O_1), verifying knowledge retention directly after the learning videos. We present the email assessment task results (O_2) in Section 4.3. Finally, we conclude by analyzing the actual behavior of our participants with the phishing study (O_M) in Section 4.4.

4.1 Pretest Analysis

In the pretest, learners achieved a mean score of 6.07 points (*Standard Deviation*, $\sigma = 2.73$) out of 9 points available. Figure 2 presents an overview of the results divided by the later consumed treatment. All three groups show very similar performances, with medians almost identical at 6.5 points for *Slides* and *Interview* and 7.0 points for participants in the *Animation* group. A Kruskal-Wallis test rejects the hypothesis of statistically significant differences between the groups ($p = 0.86; p > 0.05$), thus proving no significant difference between them.

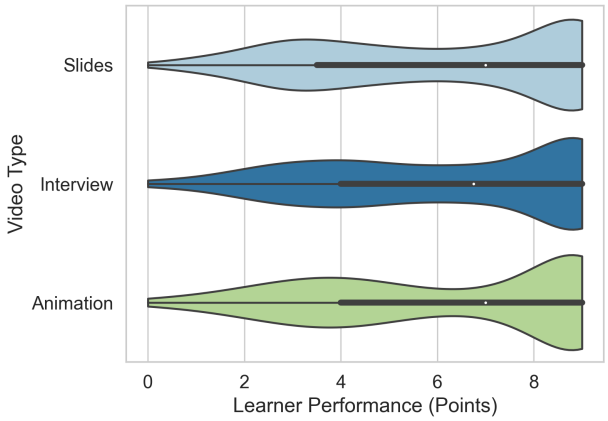


Fig. 2: Overview of learner performance in the pretest (O_P). N=1,502

4.2 Perception: Results from O_1

The overall score for learners completing the knowledge retention quiz averaged 4.5 ($\sigma = 0.59$) out of 5 points (90%). Figure 3 presents violin plots of learners' performance categorized by the video type they previously consumed. The video types of *Interview* (mean: 4.42, $\sigma = 0.68$) and *Animation* (mean: 4.40, $\sigma = 0.56$) show negligible differences. However, learners who previously consumed the *Slides* video performed significantly better, with a mean result of 4.68 ($\sigma = 0.47$). A Kruskal-Wallis test proves statistical significance ($p = 2.64 * 10^{-8}; p < 0.001$). The post hoc test (Dunn Bonferroni) revealed significant differences between learners of the *Slide* group and both other video types. We observe a moderate effect size ($d_{S;A} = 0.54$) between the groups *Slides* and *Animation* ($z = 5.68, p = 4.03 * 10^{-8}$). Between learners of *Slides* and *Interview* ($z = 4.21, p = 7.54 * 10^{-5}$), we observe a weak effect ($d_{S;I} = 0.41$).

4.3 Protection: Results from O_2

For the second dimension of cybersecurity awareness — *Protection* — we asked learners to evaluate which of the four given emails should be considered phishing. The mean result for the self-test was 3.65 ($\sigma = 0.59$) out of 4 available points. Figure 4 presents the distribution of points as violin plots for the three respective video types.

Learners having watched the *Slides* video performed weakest with a mean score of 3.58 ($\sigma = 0.61$). This is followed by learners having watched the *Interview* (mean: 3.65, $\sigma = 0.64$). Learners performing best had previously watched the *Animation* video (mean: 3.72, $\sigma = 0.51$). The slight difference between the performance of the three groups shows to be significant, as verified with a Kruskal-Wallis test for significance ($p = 0.03, p < 0.05$). The post hoc test shows a significant but weak ($d_{S;A} = 0.25$) effect only between the groups *Slides*

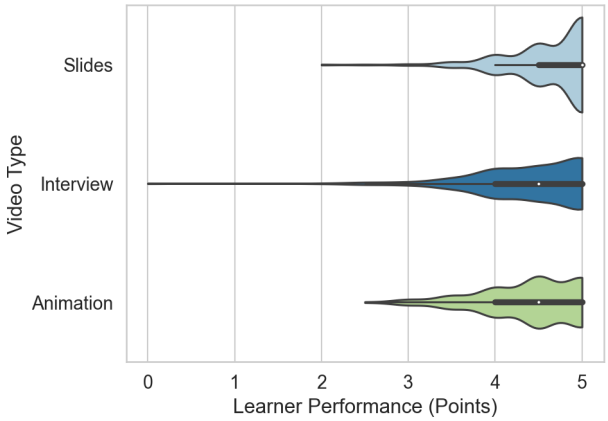


Fig. 3: Learner performance in the self-test for *Perception* (O_1). N=640

and *Interview* ($z = 2.55, p = 0.03$). For *Slides* and *Interview* ($z = 1.83, p = 0.20$) as well as *Interview* and *Animation* ($z = 0.76, p = 1.00$), the null hypothesis of significant differences in-between the data sets is rejected in the post hoc test.

4.4 Behavior: Results from O_M

The third level of security awareness — *Behavior* — describes how a user behaves in real situations. Our study featured a real-world study with laypersons, the learners from our education platform, as participants. Throughout three months after the course ended, all participants received multiple emails. The emails covered different topics and features to assess different contexts. To account for those differing topics and varying difficulties of the emails, we evaluate if and how a learner has reacted to at least one of the three emails. Reactions can range from *opening* the mail, *clicking* on a link, to *submitting* personal information.

Figure 5 provides an overview of the reactions performed by learners who had previously consumed different learning videos. With phishing emails, however, ideally, a recipient does not react at all. Hence, the lower the share of reactions, the better the respective group performed. Generally, we observe that learners who consumed the *Animation* video performed best throughout all stages of reactions to a phishing mail, while learners of the *Slides* Video performed worst. When testing for statistical significance of the differences in reactions, only the response *Mail Open* showed significance. A chi-squared (χ^2) test [5] showed a significant relationship between the consumed video type and opening at least one phishing email ($\chi^2 = 6.76, p = 0.034, p < 0.05$). Only 34% of learners who watched the *Animation* opened any mail, while we observed more than 7% higher open rates among learners consuming the *Interview* (41%) and the *Slides* Video (47%). The indicated trend of enhanced awareness among learners of the *Animation* group appears to continue in Figure 5. However, we could not

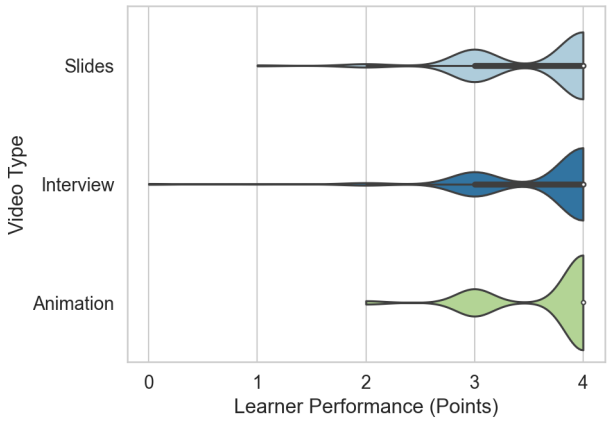


Fig. 4: Learner results of self-test for *Protection* (O_2). N=636

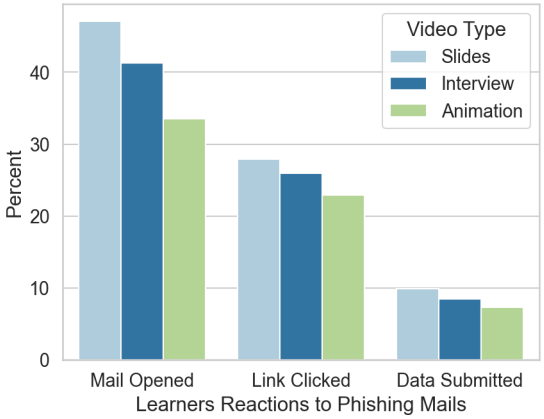


Fig. 5: Reactions of learners to the sent phishing emails (O_M), N=540

prove statistical significance for either of the two reactions *Link Click* ($\chi^2 = 1.17, p = 0.556$) or *Data Submission* ($\chi^2 = 0.77, p = 0.68$).

5 Discussion

In the previous analysis of the pretest as well as the three different tests for the three dimensions of cybersecurity awareness we introduced in Section 2.2, we observed the following results:

R1 Learners having watched the *Slides* perform best for *Knowledge*
R2 Learners having watched the *Animation* perform best when actively asked to assess emails. Thus they are best educated for the dimension of *Protection*
R3 Learners having watched the *Animation* show the best *Behavior* in real-life phishing scenarios

The observed results indicate that different education methods can help to focus the learner's experience on various levels of understanding. Concretely, in our case of cybersecurity awareness, the dimensions *Perception*, *Protection*, and *Behavior*. This finding, however, deviates from findings of previous research, such as [28,16,31] (c.f. Sec. 2.1). We therefore refrain from drawing generalizations and challenge our results and methodology in the following paragraphs.

We discuss potential alternative explanations (Sec. 5.1) for our observed results as well as threats to vailidity of our study (Sec. 5.2). We further discuss and list limitations to the results presented in this work, which in itself could be subject to further exploration in future research (Sec. 5.3).

5.1 Alternative Explanations

Our observations on the impact of the different types of educational videos derive from learners exposed to different types of videos. However, there are potential alternative explanations for the results.

Different Content in the Videos *Hypothesis: Different learning results could derive from different content being presented in the different videos.*

In principle, all three videos covered the same pieces of information on the topic of phishing. Such were cues to identify phishing emails or the motivation of cybercriminals. To do so, all videos followed a content script developed based on the animation video. While all videos covered that same information, some types of videos inherently contained more dialogue accompanying that information. One example is the *Interview*, in which the speakers covered more accompanying information than initially planned, and hence the content structure was less strict than in the other two videos. This might have led to additional content being covered, e.g., in the *Interview* in comparison to the *Slides* or *Animation*. Still, the learners of *Animation* videos performed best. Further, this insecurity on the actual content is inherent with the format (e.g. *Interview*) and should therefore be considered when selecting how content is prepared.

Knowledge Levels for Protection and Behavior *Hypothesis: The difference in the tests for the dimensions of Protection and Behavior is caused by learners having had a different skill level in advance.*

While we employed a pretest for knowledge (c.f. Section 4.1), which showed no significant difference between results for the three groups of learners, we did not employ a similar pretest for the other two dimensions. We used randomization to distribute learners equally into the three groups and hence expect that previous knowledge for all three dimensions is distributed equally among the learners. Still, we cannot claim with certainty that all groups have been equal before the treatment. In a follow-up study, we would pursue pretests for all three levels of learning and understanding.

The Factor "New" *Hypothesis: Learners are not used to animations or interviews and hence pay closer attention to the content of those.*

Our learners are used to videos in the style of *Talking Head with Slides*, as we employed in treatment X_1. A difference in learning video style could hence have caused additional interest and increased attentiveness during the videos and might have therefore helped to achieve higher learning success. Unfortunately, with our study setup, we cannot counteract this bias.

5.2 Threats to Validity

Discussing the (threats to) validity of any study helps to reflect on potential limitations for interpretation of the results and possible room for improvement in future studies. While experimental study methodologies as employed in our study are generally more robust against threats to validity [4], we want to highlight:

Threats to Internal Validity The interaction in-between criteria of internal validity is a challenge for any study. As we cross-compare learners of different types of treatments for different dimensions of cybersecurity awareness, we are relatively confident that we properly faced potential influences in-between the threats to validity such as *Maturity*, *Mortality*, or *Selection* of learners. To ensure further robustness, we compared the results of learners who did not consume any of the videos to those of the other three groups. Throughout all three tests (O_1, O_2, O_M), learners who had not watched any learning video performed worse than participants in any of the other three groups. We are hence confident that our observations are caused by the treatments X_1 through X_3.

Threats to External Validity Singular studies cannot achieve external validity interpreted as generalizability of study results. Aaronson and Ellsworth state that generalizability can only be inferred through repeated testing and verification of hypotheses across different studies in various contexts [1]. We performed our analysis in the niche of cybersecurity education, evaluating our learners' sensitization to cybersecurity threats. External validity, as generalizability, can

only be achieved through repetition of the study in different contexts, such as, e.g., misinformation on social media. The particular limitations of our study in terms of unobserved or unaccounted variables are presented in Section 5.3

5.3 Limitations and Study Group

We previously discussed potential alternative explanations to our study result as well as threats to the validity of our study results. Still, we refrain from drawing any generalizable conclusions from our study, as there are too many variables that we did not assess in our first take on the impact of video types on learning outcomes, as presented in this work. This section provides an overview of the variables not challenged in the current study and hence subject to future research. We also provide an overview of the general conditions in which the analysis was performed, such as participant demographics.

The cybersecurity course we performed this study alongside is aimed at beginners. Still, participants were rather technical, as self-supplied during a post-study survey (distribution of IT-Experience: 7.8% *Beginner*, 52.5% *Advanced*, 39.5% *Expert*). From the study population who provided their demographics in the optional post-study survey, 75,9% were male, and 23.9% self-identified as female. The shares (S) of participants who self-supplied their ages is distributed as follows: $S_{<20\ Years} = 2.0\%$; $S_{20-29} = 4.4\%$; $S_{30-39} = 7.9\%$; $S_{40-49} = 15.4\%$; $S_{50-59} = 15.4\%$; $S_{60-69} = 25.7\%$; $S_{>70\ Years} = 14.2\%$. As observable from the screenshots in Figure 1, we recorded the videos used for our study in the same professional studio as the remainder of the course. The presenters, were different from the presenter or the remainder of the MOOC. Still, all presenters were known to the participants, as they had been introduced as part of the teaching team and had presented few other previous videos. For this study and analysis, we did not investigate learning analytics information of the video, such as playback speeds, jumps between content, or quitting the video earlier, as we could not get that information for the external video. The learning success towards learners' sensitization, as observed in O_M, was measured through the user's reactions to three phishing emails. We designed the study to account for variations in difficulty between the three emails by accumulating results between all three. Therefore, however, we cannot observe changes in performance regarding time passed since the learning activity. A dedicated study would have to be performed as a follow-up.

6 Conclusion

This manuscript contributes to the body of research on the impact of different learning video styles. Previous research showed inconclusive about the effects of video styles on different types of challenges and exercises that learners are presented with. We, therefore, studied the impact of video styles on cybersecurity awareness with a cohort of more than 1,500 learners in the context of a German-speaking MOOC.

While we identified that learners of traditional *Talking Head with Slides* videos performed significantly better for recall tasks, learners exposed to *Animation* videos best applied the knowledge in different transfer tasks, such as identifying phishing emails, however, we refrained from drawing generalizable conclusions, as we are biased by the fact that both *Animation* and *Interview* videos are relatively uncommon in our MOOCs and therefore potentially particularly interesting for the learners.

We contribute more arguments and aspects to the discussion on the impact of video types, with, to the best of our knowledge, the first identification of differences in learning success for *Transfer* tasks. We aim to motivate fellow researchers to imitate our study or to perform similar ones in their contexts. For our context of cybersecurity, we welcome the fact that the shorter animation videos provide better measures towards the learners' security awareness, as they are easier to disperse, e.g., means of social media.

References

1. Aronson, E., Ellsworth, P.C.: Methods of research in social psychology. McGraw-Hill Humanities, Social Sciences & World Languages (1990)
2. Bloom, B.S.: Taxonomy of Educational Objectives: The Classification of Educational Goals. Longmans, Green (1956)
3. Bundeskriminalamt: Bundeslagebild Cybercrime 2021. Bundelagebild Cybercrime, Germany (Sep 2022), https://www.bka.de/DE/Presse/Listenseite_Pressemitteilungen/2022/Presse2022/220509_PM_CybercrimeBLB.html
4. Campbell, D.T., Stanley, J.C.: Experimental and quasi-experimental designs for research. Wadsworth, Belomt, CA (2011)
5. Chernoff, H., Lehmann, E.L.: The Use of Maximum Likelihood Estimates in χ^2 Tests for Goodness of Fit. The Annals of Mathematical Statistics **25**(3), 579–586 (Sep 1954). https://doi.org/10.1214/aoms/1177728726
6. Cisco: Cybersecurity threat trends: phishing, crypto top the list (2021), https://umbrella.cisco.com/info/2021-cyber-security-threat-trends-phishing-crypto-top-the-list
7. Cross, A., Ashok, B., Bala, S., Cutrell, E., Datha, N., Kumar, R., Kumar, V., Parthasarathy, M., Prakash, S., Rajamani, S., Sangameswaran, S., Sharma, D., Thies, W.: Online learning versus blended learning: an exploratory study. In: Proceedings of the first ACM conference on Learning@Scale. ACM, Atlanta Georgia USA (Mar 2014). https://doi.org/10.1145/2556325.2567869
8. European Union Agency for Cybersecurity.: ENISA threat landscape 2022: July 2021 to July 2022. Publications Office, LU (2022), https://data.europa.eu/doi/10.2824/764318
9. Furnell, S.: Assessing website password practices–Unchanged after fifteen years? Computers & Security **120**, 102790 (2022), iSBN: 0167-4048 Publisher: Elsevier
10. Goodyear, P., Steeples, C.: Creating shareable representations of practice. ALT-J **6**(3), 16–23 (1998), iSBN: 0968-7769 Publisher: Taylor & Francis
11. Guo, P.J., Kim, J., Rubin, R.: 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. ACM, Atlanta Georgia USA (Mar 2014). https://doi.org/10.1145/2556325.2566239

12. Hänsch, N., Benenson, Z.: Specifying IT Security Awareness. In: 2014 25th International Workshop on Database and Expert Systems Applications. pp. 326–330 (Sep 2014). https://doi.org/10.1109/DEXA.2014.71, iSSN: 2378-3915
13. Jaeger, L.: Information Security Awareness: Literature Review and Integrative Framework. In: Proceedings of the Annual Hawaii International Conference on System Sciences. IEEE Computer Society, Waikoloa Village, HA USA (Mar 2018). https://doi.org/http://hdl.handle.net/10125/50482
14. Jaeger, L., Eckhardt, A.: Eyes wide open: The role of situational information security awareness for security-related behaviour. Information Systems Journal **31**(3), 429–472 (2021). https://doi.org/10.1111/isj.12317, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/isj.12317
15. Jampen, D., Gür, G., Sutter, T., Tellenbach, B.: Don't click: towards an effective anti-phishing training. Human-centric Computing and Information Sciences **10**(1), 33 (Aug 2020). https://doi.org/10.1186/s13673-020-00237-7
16. Kizilcec, R.F., Papadopoulos, K., Sritanyaratana, L.: Showing face in video instruction: effects on information retention, visual attention, and affect. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2095–2102. ACM, Toronto Ontario Canada (Apr 2014). https://doi.org/10.1145/2556288.2557207
17. Köhler, D., Meinel, C.: The Right Tool for the Job: Overview, Comparison and Assessment of Methods for Cybersecurity Awareness Education and Verification. Preprint (Mar 2023). https://doi.org/10.13140/RG.2.2.11102.51528
18. Köhler, D., Pünter, W., Meinel, C.: Quantitatively Exploring Phishing Susceptibility in Private Contexts. Preprint, In Review (2023). https://doi.org/10.13140/RG.2.2.21865.47201
19. Lain, D., Kostiainen, K., Čapkun, S.: Phishing in Organizations: Findings from a Large-Scale and Long-Term Study. In: 2022 IEEE Symposium on Security and Privacy (SP). pp. 842–859 (May 2022). https://doi.org/10.1109/SP46214.2022.9833766, iSSN: 2375-1207
20. Lorin W. Anderson, David R. Krathwohl, Airasian, P., Bloom, B.S., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., Wittrock, M.: A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman (2001)
21. Majid, S., Kay Khine, W., Oo, M., Lwin, Z.: An Analysis of YouTube Videos for Teaching Information Literacy Skills. vol. 126, pp. 143–151 (Jan 2012)
22. McKight, P.E., Najab, J.: Kruskal-Wallis Test. In: The Corsini Encyclopedia of Psychology. John Wiley & Sons, Ltd (2010). https://doi.org/10.1002/9780470479216.corpsy0491
23. Metzger, C., Waibel, R., Henning, C., Hodel, M., Luzi, R.: Anspruchsniveau von Lernzielen und Prüfungen im kognitiven Bereich. Universität St. Gallen (1993)
24. Moore, R.L., Blackmon, S.J.: From the learner's perspective: A systematic review of MOOC learner experiences (2008–2021). Computers & Education **190**, 104596 (Dec 2022). https://doi.org/10.1016/j.compedu.2022.104596
25. Pearman, S., Zhang, S.A., Bauer, L., Christin, N., Cranor, L.F.: Why people (don't) use password managers effectively. In: Fifteenth Symposium On Usable Privacy and Security (SOUPS 2019). USENIX Association, Santa Clara, CA (2019)
26. Pearson, K.: X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling (Jul 1900). https://doi.org/10.1080/14786440009463897

27. Santos-Espino, J.M., Afonso-Suárez, M.D., Guerra-Artal, C.: Speakers and boards: A survey of instructional video styles in MOOCs. Technical Communication **63**(2), 101–115 (2016), iSBN: 0049-3155 Publisher: Society for Technical Communication
28. Steinbeck, H., Zobel, T., Meinel, C.: Using the YouTube Video Style in a MOOC: (Re-)Testing the Effect of Visual Experience in a Field-Experiment. In: Proceedings of the Ninth ACM Conference on Learning @ Scale. pp. 142–150. L@S '22, Association for Computing Machinery, New York, NY, USA (Jun 2022). https://doi.org/10.1145/3491140.3528268
29. Stolberg, H.O., Norman, G., Trop, I.: Randomized Controlled Trials. American Journal of Roentgenology **183** (2004)
30. Vogt, W.P., Johnson, R.B.: The SAGE Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences. SAGE Publications (Sep 2015)
31. Wang, J., Antonenko, P.D.: Instructor presence in instructional video: Effects on visual attention, recall, and perceived learning. Computers in Human Behavior **71**, 79–89 (Jun 2017). https://doi.org/10.1016/j.chb.2017.01.049
32. Winther, P.: Mitre att&ck technique t1566: Phishing (2022), https://attack.mitre.org/techniques/T1566/