Accepted Manuscript

Verbal redundancy in a procedural animation: On-screen labels improve retention but not behavioral performance

Bjorn B. de Koning, Charlotte M.J. van Hooijdonk, Luuk Lagerwerf

PII: S0360-1315(16)30255-X

DOI: 10.1016/j.compedu.2016.12.013

Reference: CAE 3107

To appear in: *Computers & Education*

Received Date: 15 August 2016

Revised Date: 19 December 2016

Accepted Date: 26 December 2016

Please cite this article as: de Koning B.B., van Hooijdonk C.M.J. & Lagerwerf L., Verbal redundancy in a procedural animation: On-screen labels improve retention but not behavioral performance, *Computers & Education* (2017), doi: 10.1016/j.compedu.2016.12.013.

Verbal Redundancy in a Procedural Animation: On-screen Labels Improve Retention But Not Behavioral Performance

Bjorn B. de Koning[1*], Charlotte M. J. van Hooijdonk[2] & Luuk Lagerwerf[2]

[1] Erasmus University Rotterdam

[2] Vrije Universiteit Amsterdam

Author note

Björn B. de Koning, Charlotte M. J. van Hooijdonk, and Luuk Lagerwerf

[1] Department of Psychology, Education, and Child Studies, Erasmus School of Behavioral and Social Sciences, Erasmus University Rotterdam, Rotterdam, the Netherlands, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands.

[2] Department of Language, Literature, and Communication, Faculty of Humanities, Vrije Universiteit Amsterdam, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

*Address correspondence to: Björn de Koning, Department of Psychology, Education, and Child Studies, Erasmus School of Behavioral and Social Sciences, Erasmus University Rotterdam, Rotterdam, the Netherlands, Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands. Email: b.b.dekoning@fsw.eur.nl, Phone: (+31)10-4089620.

Verbal Redundancy in a Procedural Animation: On-screen Labels Improve Retention But Not Behavioral Performance

Abstract

Multimedia learning research has shown that presenting the same words as spoken text and as written text to accompany graphical information hinders learning (i.e., redundancy effect). However, recent work showed that a "condensed" form of written text (i.e., on-screen labels) that overlaps with the spoken text, and thus is only partially redundant, can actually foster learning. This study extends this line of research by focusing on the usefulness of on-screen labels in an animation explaining a procedural task (i.e., first-aid procedure). The experiment had a 2x2x2 between-subject design (N = 129) with the factors spoken text (yes vs. no), written text (yes vs. no), and on-screen labels (yes vs. no). Learning outcomes were measured as retention accuracy and behavioral performance accuracy. Results showed that on-screen labels improved retention accuracy (but not behavioral performance accuracy) of the procedure, especially when presented together with spoken text. So, on-screen labels appear to be promising for learning from procedural animations.

Keywords: Verbal Redundancy; Instructional Animation; Procedural Task; Cognitive Load

Multimedia instructions typically contain information in graphical (e.g., picture, animation) and verbal (e.g., spoken text, subtitles) format (Mayer, 2009). Whereas the graphical information usually consists of either a diagram, picture, graph or animation, for the verbal information a critical decision involves whether this information should be presented as written and/or spoken text. Nowadays, it is increasingly being acknowledged by multimedia researchers that learning from graphics is hindered when accompanied by spoken text ànd on-screen written text that simultaneously present the same verbal information rather than by spoken text alone (e.g., Moreno & Mayer, 2002). This has become known as the (verbal) redundancy effect (Mayer, 2005). To explain the impeded learning performance arising from having to process redundant information, researchers typically turn toward the cognitive theory of multimedia learning (Mayer, 2005) and cognitive load theory (Sweller, van Merriënboer, & Paas, 1998). According to these multimedia learning theories, information is processed in working memory through visual and auditory channels, both of which are limited in the amount of information that can be processed at the same time (also see Paivio, 1986). This implies that when written text, which is identical to the spoken text, needs to be processed simultaneously with the graphical information, the visual channel likely suffers from an excessive amount of cognitive load or is overloaded. Consequently, relevant information might be missed or cannot be fully understood. More specifically, reading redundant written text is assumed to induce extraneous processing (i.e., processing that does not contribute to learning) that unnecessarily takes up working memory resources, which means that less cognitive resources are available for selecting the critical information, and organizing and integrating this information into a coherent mental representation (Mayer & Johnson, 2008). Additionally, learners have to spend cognitive resources on trying to coordinate the written text with the spoken text, which may increase extraneous processing and hence cognitive load in working memory (Sweller, 2005). The unnecessary cognitive load in working memory associated with processing identical (spoken and written) verbal information and having to integrate this information from visual and auditory sources is considered as a critical factor in explaining the redundancy effect (Kalyuga, Chandler, & Sweller, 2004). Evidently, extraneous processing may disrupt cognitive processes associated with effectively organizing the information in working memory and integrating all presented information with prior knowledge. Hence, learning is likely compromised.

Whereas the negative consequences of verbal redundancy are most likely to occur if the visually and auditory presented verbal information are identical (Adesope & Nesbit, 2012) and if the verbal information is lengthy (Kalyuga, 2012; Schüler, Scheiter, & Gerjets, 2013), recent work shows that written text that has a lower degree of similarity to the spoken text, and thus is only partially redundant, can actually foster learning. More specifically, Adesope and Nesbit (2012) found in their meta-analysis that a particular type of "condensed" verbal information (i.e., a few on-screen words summarizing the key points of the spoken text) added to a narrated multimedia presentation resulted in higher learning gains than fully redundant on-screen text. More recent studies have provided converging evidence for this finding (e.g., Roscoe, Jacovina, Harry, Russell, & McNamara, 2015; Yue, Bjork & Bjork, 2013). In other words, presenting so-called "on-screen labels" simultaneously with spoken text and graphical information is associated with improved learning, even though this constitutes partially redundant verbal information. It has been proposed that such a "reverse" redundancy effect occurs because on-screen labels emphasize the relevant terms and phrases of the verbal explanation and after having been noticed may draw learners' attention to the critical information that then can be processed more elaborately (Mayer & Johnson, 2008). Next to fulfilling such a signaling function (see Van

Gog, 2014; De Koning, Tabbers, Rikers, & Paas, 2009), another aspect that might make partially redundant verbal information (i.e., on-screen labels) effective for learning is that it is presented in the visual display at a location close to the graphical element that it refers to (cf. spatial contiguity effect; Mayer, 2009). This may increase the possibility that relevant verbal information is integrated with the graphical information which fosters learning. Together, both theoretical and empirical considerations favor the presentation of on-screen labels along with spoken text and graphical information.

So far, beneficial effects of adding on-screen labels to narrated multimedia presentations have been investigated and confirmed in a growing number of studies. These studies have focused on learning causal relations in natural and technical phenomena from sequentially (Johnson & Mayer, 2008; Yue et al., 2013) and simultaneously presented pictures (McCrudden, Hushman, & Marley, 2014), comprehension of a biological system through a self-paced, interactive diagram (Ari et al., 2014), acquiring writing strategies from a video incorporating a talking pedagogical agent (Roscoe et al., 2015), and learning foreign language vocabulary using simple and very short animated pictures representing the to-be learned words (Samur, 2012). The present study contributes to this prior work and extends it in three ways. That is, we focus on investigating the usefulness of on-screen labels (1) in an instructional animation, (2) for learning a procedural task (i.e., first-aid procedure), and (3) applying a behavioral performance measure as outcome variable in addition to a cognitive performance measure (e.g., retention test).

By focusing on on-screen labels in learning a procedural task from animation, we contribute to an emerging line of research. Typically, the topics investigated in animation research are the functioning of causal systems in biological, technical, or natural phenomena (Mayer, 2009). There is much less research on learning procedural tasks, that is, tasks involving motor actions that a person performs to achieve the required goal (Van Hooijdonk & Krahmer, 2008; for a more in-depth discussion, see Van Genuchten, van Hooijdonk, Schüler, & Scheiter, 2014). In such tasks, learners need to acquire procedural motor knowledge, i.e., knowing the object and its spatial relations and the actions that have to be performed (Glenberg, 1997). Animations are very suitable for visualizing procedural information because they can simultaneously present all of these aspects within a single display. Recently, the interest in learning procedural-motor skills from animations and/or videos has grown considerably mainly due to the widespread availability of instructional animations and videos aimed at teaching these skills on the World Wide Web and/or as offline supplementary learning materials. This has resulted in an increasing number of studies on this topic during the past few years (e.g., Castro-Alonso, Ayres, & Paas, 2015; Ganier & de Vries, 2016; Marcus, Cleary, Wong, & Ayres, 2013; Wong, Castro-Alonso, Ayres, & Paas, 2015). However, these studies are primarily interested in whether animations and videos are more suitable for teaching simple hand-manipulative tasks, like folding origami figures, than static pictures and typically involve animations and videos that are not accompanied by verbal information. Even in the studies in which verbal information is added to the animations and videos, it is done in the form of spoken text. So, it is yet unknown whether on-screen labels facilitate or hinder learning from procedural animations. For reasons mentioned earlier, it could be expected that on-screen labels help to draw attention at the right moment to the right part of the depicted procedure within the animation facilitating the integration of (visual and auditory) verbal and graphical information. Particularly for animations this might be helpful given that they continuously present information in rapid succession requiring learners to look at the right information at the right time (De Koning, Tabbers, Rikers, & Paas, 2010, 2011; De Koning & Jarodzka, in press). Moreover, for the procedural task investigated in this study (i.e., safely removing a victim from an area of danger) the verbal information is needed in order to accurately perform the procedure. Therefore, it is also possible that the on-screen labels serve as “memory pegs” on which more detailed narrated explanations regarding the procedure is hung, supporting learners in organizing (and hence retrieving) the relevant information in their mental representation.

Furthermore, the fact that we examine learning a procedural task offers the possibility to investigate whether the beneficial effects of on-screen labels can also be obtained on measurements other than those targeting learners’ understanding at the cognitive level. More specifically, we can get insight into learners’ behavioral performance by measuring how well the learned procedure is executed instead of only relying on more traditional measures that examine how well the presented information is incorporated into an accurate mental representation. By asking learners to actually perform the procedure as it would occur in a real-life situation, the learning and transfer of information can be assessed more directly and in an ecologically-valid way than having to infer this from typically used questionnaires. Until now, such a more direct testing approach has not been applied to the study of on-screen labels so it is unclear whether effects of on-screen labels translate to actual behavior. In doing so, the present study answers the call of recent proposals in aligning the type of information that should be learned (i.e., a procedural task) to the performance measure (i.e., executing the task) (Lowe, Schnotz, & Rasch, 2011).

In sum, the present study aimed to investigate whether on-screen labels are helpful in learning a procedural task from an animation. To this end, participants studied an animation explaining how to safely remove a victim from an area of danger in which we varied whether the verbal information was presented as spoken text, written text, and/or on-screen labels. Based on the above, we expected that adding on-screen labels to the animation would be particularly helpful when presented together with spoken text. So, participants studying this partially verbal redundant presentation should be able to more accurately describe (i.e., retention accuracy) and perform (i.e., behavioral performance accuracy) the learned procedure after studying the animation than when no on-screen labels are presented or when on-screen labels are presented along with the complete verbal information as written text or as written and spoken text. Furthermore, it was explored whether on-screen labels could be used as a replacement for complete verbal information without reducing retention accuracy and behavioral performance accuracy. Given that in the animation the different steps in the procedure are clearly visualized, it is possible that only a few keywords summarizing the complete verbal information are sufficient to understand the procedure.

Method

Participants and design

Participants were 129 university students (95 females, 34 males) from a large university in the north-western part of the Netherlands. Their mean age was 24.05 years (*SD* = 8.19) and 62% of them had no prior experiences with first-aid instructions. The study conformed to a 2 × 2 × 2 design with the factors Spoken Text (yes vs. no), Written Text (yes vs. no), and On-screen Labels (yes vs. no) as between-subjects factors. As depicted in Table 1, this yielded eight experimental conditions to which participants were randomly assigned. The participants in the eight conditions were comparable concerning their age ($F(7,121) = 1.09$, $p = .372$), gender ($\chi^2(7) = 7.20$, $p = .408$), and prior knowledge regarding first-aid procedures ($\chi^2(7) = 9.84$, $p = .198$). Participation was voluntarily and participants received a small monetary reward (5 euros) after the study. Participants gave informed consent based on printed information about the purpose of the study.

INSERT TABLE 1 ABOUT HERE

Materials

The materials used in this study were presented in a computer-based learning environment, which was developed in Qualtrics. The materials were electronically presented on a 22” computer screen.

Demographic questionnaire

The demographic questionnaire asked participants to indicate their gender, age, educational level, and prior knowledge concerning first-aid procedures. In line with Van Genuchten et al. (2014), prior knowledge was gauged by asking participants to indicate whether or not they had previously followed a first-aid course (yes vs. no).

Animation

The animation used in this study was obtained from the online first-aid course of the Orange Cross (www.mijnehbo.nl). The animation (see Figure 1, upper panel) showed a 12-step first-aid procedure on how to move an unconscious victim from areas of danger. The entire presentation lasted 84 seconds. In the conditions including spoken text and/or written text, the animation was accompanied by a text of 183 words which described the steps involved in the depicted first-aid procedure. The spoken text was presented in a male voice. The written text was presented as subtitles at the bottom of the computer screen just below the animation. For the conditions with on-screen labels, in line with Mayer and Johnson (2008), each step described in the text was rephrased into a two- or three-word description that summarized the key information for that step. This resulted in 12 on-screen labels that were presented one at a time when the moment the step it referred to was shown in the animation. On-screen labels were placed within the animation above the corresponding step in the first-aid procedure. In the conditions wherein participants received spoken text and/or written text and/or on-screen labels, the different types of verbal information were presented simultaneously (see Figure 1, bottom panel).

In four experimental conditions, this resulted in redundant presentation of verbal information. In the conditions wherein on-screen labels were combined with spoken text or written text there was a partial overlap between the two sources of verbal information. Complete overlap between the different sources of verbal information was present in the condition with spoken text and written text as well as in the condition with spoken text, written text, and on-screen labels. In four conditions there was no verbal redundancy: animation only, animation with spoken text, animation with written text, and animation with on-screen labels. Participants had the possibility to navigate through the animation with a start/pause button and a slider to forward/rewind the presentation. They could watch the animation as often as they liked. To control for possible between-group differences regarding their learning time, we measured participants' time spent (in seconds) on studying the animation, which was automatically logged in the computer-based learning environment.

INSERT FIGURE 1 ABOUT HERE

Learning outcomes

To test how much knowledge participants had obtained from studying the animation, we used a behavioral performance test and a retention test. The behavioral performance test asked participants to execute the first-aid procedure they had studied using a first-aid dummy. For this task, a first-aid dummy was lying on the floor and participants had to move it in a way they had learned by pretending the dummy was a victim that had to be moved from an area of danger. Participants were not provided any feedback on their performance during or after this task. This way, it was possible to straightforwardly measure participants' procedural motor knowledge regarding the first-aid procedure. Participants' behavioral performance was videotaped from two angles to enable scoring afterwards. Behavioral performance accuracy was measured by the number of steps that were performed both correctly and in the correct order. Participants could score either 0 or 1 points per step. To assess the interrater reliability for behavioral performance accuracy, two raters independently coded 20% of all videos. The Cohen's kappa was .76. The remaining 80% of the data were scored by a single rater only. The total number of correctly performed steps was calculated resulting in a maximum of 12 points. This score was converted into a percentage reflecting the behavioral performance accuracy.

The retention test was used to measure the amount of information that participants could remember from the instruction. It consisted of a cloze-test that contained the entire 183-word text from the instruction in which 12 words were missing. Participants' task was to fill in the blanks with the correct words based on what they had learned during studying the animation. Participants could score either 0 or 1 points per answer. To assess the interrater reliability on the retention test, two raters independently coded 20% of all answers. Cohen's kappa was .93. Based on this, it was decided that the remaining 80% of the data was scored by a single rater. The total number of correctly performed steps was calculated resulting in a maximum of 12 points. This score was converted into a percentage reflecting their accuracy in describing the first-aid procedure (i.e. retention accuracy).

Procedure

Participants were tested individually in a single session of approximately 30 minutes. First, they read and signed the informed consent form. Next, they familiarized themselves with performing a simple first-aid procedure on a first-aid dummy. For this task, which targeted a first-aid procedure unrelated to the one in the learning phase, participants received a paper-based instruction describing three steps on how to tilt a victim onto its side and back. Subsequently, participants were seated in front of a computer screen and filled in the demographic questionnaire. After that, participants were instructed to study the animation and were informed that they would be tested for their knowledge after this study phase. Depending on their condition, they watched the animation with/without spoken text, with/without written text, and with/without on-screen labels. After studying the animation, participants engaged in a distraction task in the form of a Sudoku puzzle which they had to complete within 10 minutes. Participants then completed the behavioral performance test followed by the retention test. No time limits were set for the behavioral performance test and the retention test. Finally, participants were thanked for their participation and received their monetary reward.

Results

A 2 × 2 × 2 multivariate analysis of variance (MANOVA) with the between-subjects factors Spoken Text (yes vs. no), Written Text (yes vs. no), and On-screen labels (yes vs. no) was conducted on the behavioral performance accuracy and retention accuracy scores. Learning time was not included as a covariate in this analysis given that the time spent on studying the animation did not differ statistically between the eight experimental conditions ($F(7,121) = 1.49$, $p = .178$).

INSERT TABLE 2 ABOUT HERE

The MANOVA analysis revealed a significant main effect of Spoken Text (*Wilks's* $\lambda = .85$, $F(2,120) = 10.86$, $p < .001$, $\eta_p^2 = .15$). Overall participants in the conditions with spoken text showed higher retention accuracy and behavioral performance accuracy regarding the first-aid procedure than participants in the conditions without spoken text (retention accuracy: $F(1,121) = 21.69$, $p < .001$, $\eta_p^2 = .15$; behavioral performance accuracy: $F(1,121) = 5.77$, $p = .018$, $\eta_p^2 = .05$). There was also a significant main effect of Written Text (*Wilks's* $\lambda = .83$, $F(2,120) = 12.42$, $p < .001$, $\eta_p^2 = .17$), showing that overall participants in the conditions with written text had higher retention accuracy regarding the first-aid procedure than participants in the conditions without written text ($F(1,121) = 23.43$, $p < .001$, $\eta_p^2 = .16$). No such significant differences were found on behavioral performance accuracy ($F < 1$). The main effect for On-screen Labels appeared not to be significant (*Wilks's* $\lambda = .98$, $F(2,120) = 1.11$, $p = .332$, $\eta_p^2 = .02$).

Regarding the interacting effects, overall there was a significant two-way interaction between the factors Spoken Text and Written Text (*Wilks's* λ = .92, $F(2,120) = 5.32$, $p = .006$, $\eta_p^2 = .08$), which was observed for both retention accuracy ($F(1, 121) = 9.60$, $p = .002$, $\eta_p^2 = .07$) and behavioral performance accuracy ($F(1, 121) = 5.22$, $p = .024$, $\eta_p^2 = .04$). Bonferroni adjusted pairwise comparisons involving the four groups resulting from combining the factors Spoken Text and Written Text -thus collapsed over the on-screen label conditions- (groups: written text, spoken text, written text and spoken text, no written text and no spoken text) showed the following results (see Table 2). Participants in the written text group (NST-WT-L and NST-WT-NL together) had higher retention accuracy ($p < .001$) regarding the first-aid procedure than participants in the group without spoken and written text (NST-NWT-L and NST-NWT-NL together). Also, participants in the group with spoken text (ST-NWT-L and ST-NWT-NL together) obtained higher retention accuracy ($p < .001$) than participants in the group without spoken and written text (NST-NWT-L and NST-NWT-NL together). Similarly, participants in the group with spoken text and written text (ST-WT-L and ST-WT-NL together) obtained higher retention accuracy ($p < .001$) than participants in the group without spoken and written text (i.e., NST-NWT-L and NST-NWT-NL together). Furthermore, participants in the spoken text group (i.e., ST-NWT-L and ST-NWT-NL together) outperformed participants in the written text group (i.e., NST-WT-L and NST-WT-NL together) on behavioral performance accuracy ($p < .001$). No other significant differences were found.

Most interestingly, overall there was a significant three-way interaction between Spoken Text, Written Text, and On-screen Labels (*Wilks's* λ = .93, $F(2, 120) = 4.51$, $p = .013$, $\eta_p^2 = .07$), which appeared to be confined to retention accuracy ($F(1, 121) = 3.78$, $p = .054$, $\eta_p^2 = .30$; behavioral performance accuracy: $F(1,121) = 1.56$, $p = .214$, $\eta_p^2 = .01$). Bonferroni adjusted pairwise comparisons involving the eight experimental conditions showed the following results (see Table 3). Looking specifically at the extent to which on-screen labels influenced retention accuracy, it appeared that participants in the condition with spoken text and on-screen labels (ST-NWT-L) had higher retention accuracy than participants in the condition without spoken text, without written text, and without on-screen labels (NST-NWT-NL) ($p < .001$). Likewise, participants who received the animation in the condition with on-screen labels and written text (NST-WT-L) showed higher retention accuracy than participants in the condition without spoken text, without written text, and without on-screen labels (NST-NWT-NL) ($p < .001$). Presenting on-screen labels together with spoken text and written text (ST-WT-L) also resulted in higher retention accuracy than in the condition without spoken text, without written text, and without on-screen labels ($p < .001$). Participants in the condition with only on-screen labels (NST-NWT-L) showed a similar trend for higher retention accuracy than participants in the condition without spoken text, without written text, and without on-screen labels (NST-NWT-NL) ($p = .091$). In the conditions without on-screen labels that involved spoken text (ST-NWT-NL), written text (NST-WT-NL) or both (ST-WT-NL) higher retention accuracy was obtained than in the condition without spoken text, without written text, and without on-screen labels (NST-NWT-NL) ($p < .001$). Participants in the condition with spoken text and written text (ST-WT-NL) also obtained higher retention accuracy than participants in the condition with on-screen labels (NST-NWT-L) ($p = .019$). No other significant differences were found.

INSERT TABLE 3 ABOUT HERE

Discussion

This study aimed to investigate whether previously reported beneficial effects of adding short on-screen text (i.e., on-screen labels) to multimedia instructions, referred to as the "reverse" redundancy effect (e.g., Mayer & Johnson, 2008), also apply to learning a procedural task from an animation. Consistent with prior research, we found that presenting on-screen labels in a procedural animation resulted in improved retention, but not actual behavioral performance, of the studied procedure (e.g., Mayer & Johnson, 2008; McCrudden et al., 2015).

Specifically, we found that presenting the animation with on-screen labels and the complete verbal information in the form of spoken text (ST-NWT-L), written text (NST-WT-L) or both (ST-WT-L) resulted in higher retention outcomes compared to the animation without any verbal information (NST-NWT-NL). Results were in the same direction for learners who studied the animation with only on-screen labels (NST-NWT-L) although this effect was less pronounced. This suggests that in learning from a procedural animation, on-screen labels alone offer too limited verbal information to complement the graphic information when no other verbal information is provided. Especially when presented together with the complete spoken and/or written text on-screen labels appear to be help learners remember the information shown in a procedural animation. This is consistent with predictions derived from relevant theoretical models, like the cognitive theory of multimedia learning (Mayer, 2005) and the cognitive load theory (Sweller et al., 1998). According to these models, on-screen labels are assumed to facilitate learners' cognitive processing by helping them to identify and select relevant verbal and graphical information and establishing a link between them so as to better organize it in memory. In other words, on-screen labels seem to reduce extraneous cognitive processing associated with searching for and connecting relevant pieces of information, leaving cognitive resources available for constructing a coherent mental representation of the multimodal information (Sweller et al., 1998).

Together, our findings support prior research on on-screen labels in multimedia learning, both theoretically and empirically, showing that presenting redundant information, in the form of short on-screen text presented in close proximity of the associated pictorial information, may actually foster rather than impede learners' cognitive processing of multimedia materials. Importantly, we extend prior findings by demonstrating that this "reverse" redundancy principle also applies to multimedia materials that are transient in nature (i.e., animation) and thus are not constantly available to learners. Evidently, helping learners to attend to the relevant information at the right time is particularly relevant in animations and other sorts of dynamic visualizations (e.g., De Koning & Jarodzka, in press; De Koning, Tabbers, Rikers, & Paas, 2011). On-screen labels seem to provide a useful means to accomplish this and may provide an alternative or complement to other types of attention guidance such as color coding to improve processing of multimedia instructions (Van Gog, 2014). Moreover, this study extended previous multimedia research by focusing on learning a procedural task (i.e., first-aid procedure) as opposed to a more causal learning task (e.g., lightning formation). Studies on learning from multimedia instructions, especially those involving on-screen labels, typically require learners to study causal chains of events in natural, technical or biological systems. In the present study, we broaden the scope of this research by addressing a topic that has, in comparison, received much less attention, namely procedural learning. Based on the above issues, our study contributes to a better understanding of the generalizability of the "reverse" redundancy principle in learning from multimedia instructions.

Another contribution of the present study is that it examined the effects of on-screen labels at the behavioral level. That is, we asked learners to perform the learned procedure and looked at how well they performed that procedure. So far, the effects of on-screen labels have only been examined at the cognitive level by looking at how well learners were able to memorize the presented information (as discussed above) and/or the extent to which they were able to reason about the learned information using a verbal transfer test. Prior research generally does not show positive (nor negative) effects of on-screen labels on tests tapping into higher-order cognitive processing (e.g., reasoning) such as transfer tests (Mayer & Johnson, 2008; McCrudden et al., 2015; for an exception see Yue et al., 2013). Consistent with this, our findings indicate that learning the first-aid procedure from an animation including on-screen labels (either with –ST-NWT-L and NST-WT-L and ST-WT-L– or without –NST-NWT-L– complete verbal information) did not help learners to better perform the procedure than learning without on-screen labels (ST-NWT-NL; NST-WT-NL; ST-WT-NL; NST-NWT-NL). Possibly, learners primarily relied on the complete verbal information (and the movements in the animation) to perform the learned procedure. This verbal and pictorial information likely provided more guidance than the on-screen labels in how to exactly perform the different activities in the procedure. For example, one of the keywords summarized the verbal explanation as “seat victim upright” while this did not indicate that the movement should be performed fluently. Hence, the on-screen labels merely helped to remember the different activities, but were presumably less helpful to support learners in actually performing the procedure as accurate as possible.

Together, our findings combined with those obtained in prior research suggest that beneficial effects of on-screen labels do not easily translate to measures that move beyond measuring the amount of information that learners remember from multimedia instructions (e.g., retention test). Given that this has now been demonstrated with multimedia instructions wherein static pictures (Mayer & Johnson, 2008), animations (present study), and informational text (McCrudden et al., 2015) form the primary source for presenting the information, it seems useful to start investigating why this is the case and attempt to find ways to improve performance on measures other than those tapping into memorized information. For example, it could be examined whether on-screen labels are more effective to facilitate behavioral performance of a procedure if the type of information summarized by the on-screen label focuses on subtleties on how to perform the movement instead of simply providing a mnemonic for the activity more generally in the respective step.

Limitations and future research

A drawback of this study is that with eight experimental conditions the total number of participants was relatively low and preferably should have involved more participants per condition. It is possible that in the present study the low retention accuracy of the condition studying the animation without spoken text, without written text, and without on-screen labels (NST-NWT-NL) is due to the relatively low sample size and in this way contaminates the results. It is however also possible that this low score provides an accurate picture, which would be in line with our findings mentioned earlier suggesting that at least some type of verbal information (complete text and/or labels) is needed to remember the steps in the first-aid procedure. In a similar vein, unexpectedly the highest retention accuracy was obtained in the condition with both spoken text and written text, without on-screen labels (ST-WT-NL). It is unclear how to this result has come about; a likely possibility is that this high score is related to the low sample size. Overall, given the low sample size the results of this study need to be interpreted with caution. Future studies replicating this study with more (e.g., at least 20) participants per condition are needed to provide more solid findings enabling stronger conclusions about the usefulness of on-screen labels in procedural animations.

Furthermore, an aspect that was not addressed in this study and which requires further investigation concerns the attentional and cognitive processing during studying procedural animations, or multimedia instructions in general, with on-screen labels. As in previous studies we did not apply online process-related measures, such as eye-tracking and verbal protocols, so it is yet unclear how on-screen labels exactly operate in accomplishing their effect. Critical questions in this regard are for example how much attention is paid to the on-screen labels and/or other elements of the multimedia instruction, and to what extent learners engage in meaningful constructive processing to integrate all presented information with each other as well as with prior knowledge. Presumably, this may also provide additional insight into why effects of on-screen labels did not translate to behavioral performance accuracy (or more broadly, to other types of measures than retention/memorizing information like transfer). Regarding this latter aspect, it is important to note that a potential limitation of this study is that our animation only lasted approximately one and a half minute, which is a relatively short learning session given that in educational settings the time that is usually spend on learning new information is much longer. So, it might be worth exploring whether or not a more prolonged animation might help learners to better integrate and/or internalize the information enabling them to more accurately perform the studied procedure. Moreover, this would improve the ecological validity of the results as it would more closely resemble a realistic educational setting. Furthermore, several more generic issues remain to be addressed. These involve, among others, to what extent learner-control has influenced our results, whether similar findings will be obtained when studying an animation depicting a causal system (e.g., showing hoe a toilet tank works), and how on-screen labels interact with other design manipulations to the graphical information like visual cueing (or signaling) that are intended to support learners' selection of relevant information from the animation. Addressing these issues in future research advances our understanding on (more) effective uses of on-screen labels to improve learning from multimedia instructions.

Conclusion

This study extends research on the 'reverse' verbal redundancy effect by demonstrating the usefulness of on-screen labels in a procedural animation. This reinforces the suggestion put forward previously by other researchers (e.g., Mayer & Johnson, 2008; Yue et al., 2013) that teachers and instructional designers could improve learners' understanding of multimedia instructions if they embed short text segments (i.e., on-screen labels that are not identical to the complete verbal information) into the visual display. Our study shows that this also holds for teaching procedural knowledge (as opposed to conceptual knowledge) and when presenting the information as a learner-controllable animation, indicating that on-screen labels can be applied across various types of multimedia instructions. It should be stressed however that beneficial effects of on-screen labels are most likely to occur on measures that tap into learners' memory for the studied information. This means that on-screen labels should primarily be used in the context of learning goals that are aimed at helping learners to memorize information from multimedia instructions.

References

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology, 104(1)*, 250-263. doi:10.1037/a0026147

Ari, F., Flores, R., Inan, F. A., Cheon, J., Crooks, S. M., Paniukov, D., & Kurucay, M. (2014). The effects of verbally redundant information on student learning: An instance of reverse redundancy. *Computers & Education*, *76*, 199-204. doi: http://dx.doi.org/10.1016/j.compedu.2014.04.002

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2015). Animations showing Lego manipulative tasks: Three potential moderators of effectiveness. *Computers & Education, 85*, 1-13. doi: 10.1016/j.compedu.2014.12.022

De Koning, B. B., & Jarodzka, H. (in press). Attention guidance strategies for supporting learning from dynamic visualizations. In R. Lowe and R. Ploetzner (Eds.), *Learning from dynamic visualizations: Innovations in research and practice*.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2011). Improved effectiveness of cueing by self-explanations when learning from a complex instructional animation. *Applied Cognitive Psychology*, *25*, 183-194. doi: 10.1002/acp.1661

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*, 111-122. doi:10.1016/j.learninstruc.2009.02.010

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*, 113-140. doi: 10.1007/s10648-009-9098-7

Ganier, F, & de Vries, P. (2016). Are instructions in video format always better than photographs when learning manual techniques? The case of learning how to do sutures. *Learning and Instruction, 44*, 87-96. http://dx.doi.org/10.1016/j.learninstruc.2016.03.004

Glenberg, A. M. (1997). What memory is for: Creating meaning in the service of action. *Behavioral and Brain Sciences*, *20*, 41-50. doi: 10.1017/S0140525X97470012

Kalyuga, S. (2012). Instructional benefits of spoken words: A review of cognitive load factors. *Educational Research Review, 7*, 145-159. doi:10.1016/j.edurev.2011.12.002

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46(3)*, 567-581. doi: 10.1518/hfes.46.3.567.50405

Lowe, R., Schnotz, W., & Rasch, T. (2011). Aligning affordances of graphics with learning task requirements. *Applied Cognitive Psychology, 25 (3)*, 452-459. doi: 10.1002/acp.1712

Marcus, N., Cleary, B., Wong, A., & Ayres, P. (2013). Should hand actions be observed when learning hand motor skills from instructional animations? *Computers in Human Behavior, 29 (6)*, 2172 – 2178. doi: http://dx.doi.org/10.1016/j.chb.2013.04.035

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology, 100(2)*, 380-386. doi: 10.1037/0022-0663.100.2.380

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York, NY: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd edn). New York: Cambridge University Press.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94(1)*, 156-163. doi: 10.1037//0022-0663.94.1.156

McCrudden, M. T., Hushman, C. J., & Marley, S. C. (2014). Exploring the boundary conditions of the redundancy principle. *The Journal of Experimental Education, 82 (4)* ,537-554. doi: 10.1080/00220973.2013.813368

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.

Roscoe, R. D., Jacovina, M. E., Harry, D., Russell, D. G., & McNamara, D. S. (2015). Partial verbal redundancy in multimedia presentations for writing strategy instruction. *Applied Cognitive Psychology, 29,* 669-679. doi: 10.1002/acp.3149

Samur, Y. (2012). Redundancy effect on retention of vocabulary words using multimedia presentation. *British Journal of Educational Technology, 43(6)*, 166-170. doi:10.1111/j.1467-8535.2012.01320.x

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior, 29(4)*, 1590-1601. doi:10.1016/j.chb.2013.01.047

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 159–168). New York, NY: Cambridge University Press.

Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10(3)*, 251-296. doi: 10.1023/A:1022193728205

Van Genuchten, E., van Hooijdonk, C., Schüler, A., & Scheiter, K. (2014). The role of working memory when 'learning how' with multimedia learning material. *Applied Cognitive Psychology, 28 (3),* 327-335. doi:10.1002/acp.2998

van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R.E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd rev. ed.) (pp. 263-278). New York: Cambridge University Press.

van Hooijdonk, C., & Krahmer, E. (2008). Information modalities for procedural instructions: The influence of text, pictures, and film clips on learning and executing RSI exercises. *IEEE Transactions on Professional Communication, 51,* 50–62. doi: 10.1109/TPC.2007.2000054

Wong, M., Castro-Alonso, J. C., Ayres, P., & Paas F. (2015). Gender effects when learning manipulative tasks from instructional animations and static presentations. *Educational Technology and Society, 18 (4)*, 37-52.

Yue, C. L., Bjork, E. L., & Bjork, R. A. (2013). Reducing verbal redundancy in multimedia learning: An undesired desirable difficulty? *Journal of Educational Psychology, 105*, 266-277. doi: 10.1037/a0031971

Table 1. Overview and Characteristics of the Eight Experimental Conditions

	Spoken text (ST)				No spoken text (NST)			
	Written text (WT)		No written text (NWT)		Written text (WT)		No written text (NWT)	
	Labels (L)	No labels (NL)	Labels (L)	No labels (NL)	Labels (L)	No labels (NL)	Labels (L)	No labels (NL)
Condition	ST WT L	ST WT NL	ST NWT L	ST NWT NL	NST WT L	NST WT NL	NST NWT L	NST NWT NL
N	16	16	16	16	16	16	17	16
Age*	25.75 (10.28)	24.88 (10.63)	28.00 (12.20)	23.63 (3.36)	21.25 (3.36)	22.38 (2.94)	23.82 (11.11)	22.69 (2.87)
Gender (male/female)	5/11	3/13	5/11	7/9	5/11	2/14	2/15	5/11
Prior knowledge (yes/no)	3/13	8/8	5/11	9/7	5/11	4/12	6/11	9/7

* Mean age in years followed by standard deviation (in parentheses)

Table 2. Means and Standard Deviations for Retention Accuracy and Behavioral Performance Accuracy as a Function of Spoken Text and Written Text

	Retention accuracy		Behavioral performance accuracy	
	M	*SD*	*M*	*SD*
Written text	66.19	19.07	79.83	13.07
Spoken text	65.62	16.28	83.33	12.53
Written and Spoken text	71.02	16.07	80.12	16.63
No written text and No spoken text	41.87	20.32	72.27	11.76

Table 3. Means and Standard Deviations (in Parentheses) for Retention Accuracy, Behavioral Performance Accuracy, and Learning Time as a Function of Spoken Text, Written Text, and On-Screen Labels

	Spoken text				No spoken text			
	Written text		No written text		Written text		No written text	
	Labels	No labels	Labels	No labels	Labels	No labels	Labels	No labels
Retention accuracy	69.89 (15.11)	72.16 (17.40)	65.34 (11.60)	65.91 (20.33)	62.50 (23.44)	69.89 (13.16)	50.80 (18.76)	32.39 (17.87)
Behavioral performance accuracy	78.03 (18.83)	82.20 (14.42)	87.50 (8.49)	79.17 (14.67)	82.20 (13.18)	77.46 (12.95)	74.87 (13.41)	69.51 (9.35)
Learning Time (sec.)	219.82 (63.67)	217.05 (97.67)	176.36 (64.78)	179.30 (91.28)	187.39 (53.56)	217.04 (56.42)	182.50 (114.40)	154.19 (56.63)

Figure Captions

Figure 1. Screenshot of the animation illustrating the step “Place your hands around this forearm with interconnected fingers and thumbs” (upper panel: animation without verbal information; bottom panel: animation with written text and on-screen labels).

Figure 1


Handen over onderarm
Leg uw handen, met aaneengesloten vingers
en duimen, over deze onderarm.

Highlights

- Recent research shows that on-screen labels improve learning of multimedia material
- This study aimed to extend this to learning from procedural animations
- We also applied a behavioral performance measure next to a cognitive outcome measure
- On-screen labels improved retention but not behavioral performance of the procedure
- On-screen labels appear to be promising for learning from procedural animations