Applied Cognitive Psychology, *Appl. Cognit. Psychol.* **28**: 327–335 (2014)
Published online 6 January 2014 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/acp.2998

The Role of Working Memory when 'Learning How' with Multimedia Learning Material

ERLIJN VAN GENUCHTEN[1], CHARLOTTE VAN HOOIJDONK[2]*, ANNE SCHÜLER[1] and KATHARINA SCHEITER[1]

[1]*Knowledge Media Research Center, Tubingen, Germany*
[2]*Department of Language and Communication, VU University Amsterdam, Amsterdam, The Netherlands*

Summary: The aim of the reported experiment was to obtain insight into how learners' visuo-spatial working memory is involved during learning how to perform procedural-motor tasks from a multimedia instruction (i.e. 'learning how'). Eighty-two participants studied first-aid procedures using text only or multimedia. Working memory involvement was gauged by measuring the interference between learning first-aid procedures and performing a spatial dual task. Learning outcomes were measured as task performance and task description. The results showed that performing a spatial dual task interfered to a larger extent with learning from text only than from multimedia. The results tend to support the assumption that pictures in tasks focusing on 'learning how' are beneficial to learning, because they might omit the need to engage in imagery and therewith reduce the cognitive effort that is required to understand the learning material. Copyright © 2013 John Wiley & Sons, Ltd.

INTRODUCTION

An established effect in the field of learning and instruction is the multimedia effect, which states that learning is fostered by adding pictures to text (Mayer, 2009). Most studies on the multimedia effect investigated the beneficial effect of pictures concerning the functioning of a system (i.e. causal tasks; see Mayer, 2005, 2009). Accordingly, we have fairly good knowledge about the cognitive processes associated with processing text and pictures when learning about causal tasks (Mayer, 2009; Schüler, Scheiter, & van Genuchten, 2011). On the other side, considerable fewer studies have focussed on whether pictures aid learning concerning how to do something (i.e. procedural-motor tasks), and hence, also less is known about the cognitive processes associated with processing text and pictures when learning about procedural-motor tasks. As outlined in the next section, causal and procedural tasks are characterised by different features and—on the basis of different lines of research (e.g. Fischer & Zwaan, 2008; Fogassi & Ferrari, 2011; Kosslyn, Ganis, & Thompson, 2001; Postma & Barsalou, 2009; Rizzolatti & Craighero, 2004)—it can be assumed that causal and procedural tasks differ especially regarding their affordance for mental imagery: Whereas causal tasks have a rather low affordance for imagery, the affordance for imagery with procedural tasks is rather high (for a more thorough explanation of that argument, see next section). As a consequence, it may be assumed that the multimedia effect observed for procedural tasks (e.g. Michas & Berry, 2000; Schwan & Riempp, 2004; van Genuchten, Scheiter, & Schüler, 2012; Wan & Baragash, 2011) can be traced back to different cognitive processes in working memory than the multimedia effect observed for causal tasks. Therefore, in the current paper, we investigated whether the affordance for mental imagery in procedural task has consequences for the processing of learning material that consists of text only, or text and pictures in working memory. We focused on the role of one specific working memory subsystem, namely the visuo-spatial sketchpad (Baddeley, 1999), because mental images are assumed to be processed in this part of working memory (e.g. Baddeley & Andrade, 2000; Baddeley, Grant, Wight, & Thomson, 1973).

*Correspondence to: Charlotte van Hooijdonk, Department of Language and Communication, Faculty of Arts, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.
E-mail: c.m.j.van.hooijdonk@vu.nl

'LEARNING THAT' AND 'LEARNING HOW'

In tasks that focus on the functioning of a system (i.e. causal tasks; Mayer, 2005, 2009) changes in some part of the system and their effects on other parts of the system are described and depicted (e.g. the functioning of a bicycle tire pump; Mayer & Gallini, 1990). The aim of these tasks is to 'learn that' (Ryle, 2000), that is, to obtain declarative, explicit knowledge and create a mental representation of the system's components and their causal relationships. Here, learning outcomes can be gauged by measuring how much information can be recalled regarding the structure and functioning of the system, and whether this information can be used to reason about similar situations (i.e. transfer of information; Mayer, 2009).

In contrast, in tasks that focus on learning how to do something (i.e. procedural-motor tasks; e.g. Dechsri, Jones, & Heikkinen, 1997; Iserbyt, Mols, Elen, & Behets, 2012; Michas & Berry, 2000; Van Hooijdonk & Krahmer, 2008), the motor actions that have to be performed by a person to achieve a goal are described (e.g. how to tie a knot: Hayes & Henk, 1986). The aim of these tasks is to 'learn how' (Ryle, 2000), that is, to obtain procedural-motor, implicit knowledge and develop the skill in a way that allows performing the procedure automatically and error free (cf. Millar, Lubben, Gott, & Duggan, 1994). This means that 'learning how' involves not only knowledge concerning objects and their spatial relationships (i.e. declarative/explicit knowledge) but also knowledge concerning the motor actions that have to be applied to these objects in order to achieve the goal (i.e. procedural-motor/implicit knowledge; Glenberg, 1997; Glenberg & Langston, 1992; Langston, Kramer & Glenberg, 1998). Therefore, learning outcomes can be gauged both by measuring how well declarative/explicit knowledge is recalled and by measuring how well (i.e. how correctly/efficiently/successfully) the motor actions are performed (Ryle, 2000).

Thus, in contrast to tasks that focus on 'learning that', tasks that focus on 'learning how' convey not only declarative/explicit knowledge but also information about the motor actions that have to be applied in order to achieve the goal. Interestingly, on the basis of findings from different lines of research, it can be assumed that tasks describing motor actions may have a high affordance for mental imagery (e.g. Fischer & Zwaan, 2008; Fogassi & Ferrari, 2011; Kosslyn et al., 2001; Postma & Barsalou, 2009; Rizzolatti & Craighero, 2004). For instance, it has been shown that the use of action-related words (i.e. nouns, verbs and adjectives)—as often used in tasks describing procedural-motor information—seems to trigger mental simulation (Fischer & Zwaan, 2008). According to Postma and Barsalou (2009), this mental simulation mechanism is essentially the same as mental imagery. The difference between the two mechanisms relies in the fact that mental imagery is typically assumed to be conscious, whereas mental simulation can be both, conscious and unconscious (Postma & Barsalou, 2009). Moreover, the depiction of an interaction between a body part and an object (e.g. a picture of a hand grabbing a glass) seems to automatically trigger the activation of mirror neurons in the motor system (Fogassi & Ferrari, 2011; Rizzolatti & Craighero, 2004; see also Van Gog, Paas, Marcus, Ayres, & Sweller, 2009; Marcus, Cleary, Wong, & Ayres, 2013), which has been assumed to be involved in conscious motor imagery (Kosslyn et al., 2001). Thus, tasks focusing on 'learning how' (i.e. tasks conveying procedural-motor information) may have a higher probability to afford imagery than tasks focusing on 'learning that', as in these latter tasks merely objects and their interrelations are described—without using action-related words or describing a person interacting with the system (cf. Rizzolatti & Craighero, 2004).

The question arises whether the affordance for mental imagery in tasks focusing on 'learning how' has consequences for the processing of learning material that consists of text only or consists of text and pictures, because one may assume that presenting pictures omits the need to engage in imagery (Marcus, Cooper, & Sweller, 1996). To address this question, we investigated information processing during learning by assessing the involvement of working memory during 'learning how', by using the secondary task paradigm. This paradigm is introduced in the next section.

INVESTIGATING WORKING MEMORY INVOLVEMENT DURING MULTIMEDIA LEARNING

According to Baddeley (1999), working memory consists of a central executive, which coordinates two slave systems, namely the phonological loop and the visuo-spatial sketchpad. The phonological loop is assumed to process verbal information, whereas the visuo-spatial sketchpad is assumed to be involved in processing visuo-spatial information (e.g. mental images and pictures) as well as movement control (Baddeley, 1999). As noted earlier, it can be assumed that tasks that focus on 'learning how' differ from tasks that focus on 'learning that' regarding the mental images that are triggered (Baddeley, 2012; Smyth & Pendleton, 1990). Therefore, cognitive processing between both task types should especially differ with regard to the involvement of the visuo-spatial sketchpad, where mental images are assumed to be processed (e.g. Baddeley & Andrade, 2000; Baddeley et al., 1973).

In general, to investigate information processing in working memory during learning, the dual-task approach can be used (e.g. Andrade, 2001). This approach allows measuring the load of working memory subsystems when performing the primary task, in this case multimedia learning of a procedural task. The underlying principle of performing the dual task is that it requires the resources of one of the working memory subsystems and therefore causes interference with the learning task (Andrade, 2001). If learning outcomes are affected by performing this dual task, it can be concluded that the working memory subsystem that was loaded by the dual task was involved during learning (Baddeley, 1999). The involvement of the visuo-spatial sketchpad during learning can be measured using spatial motor tasks, where participants have to continuously conduct specific movements (Farmer, Berman, & Fletcher, 1986). Because the visuo-spatial sketchpad controls the execution of movements (e.g. Lawrence, Myerson, Oonk, & Abrams, 2001; Logie & Marchetti, 1991), spatial motor tasks interfere with the processing of visuo-spatial information in this part of working memory. Usually, participants press buttons on a keyboard with their fingers (e.g. Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999; Farmer et al., 1986); however, in some studies, participants tapped with their foot instead (Liefooghe, Vandierendonck, Muyllaert, Verbruggen, & Vanneste, 2005; Miyake, Emerson, Padilla, & Ahn, 2004; Service & Turpeinen, 2001; see Schüler et al., 2011, for a review). In line with the assumption that movements in general disrupt visuo-spatial processing, Miyake et al. (2004) demonstrated that the foot-tapping task interfered with the processing of visuo-spatial information, such as colours and shapes.

In sum, stronger interference between the tapping task and information processing implies stronger involvement of the visuo-spatial sketchpad during learning. Such interference can, for example, be measured by comparing learning outcomes of participants who learned with dual task and participants who learned without dual task. If there is interference, participants in the dual-task condition will have lower learning outcomes than participants who learned without dual task.

Previous research investigating the involvement of the visuo-spatial sketchpad during multimedia learning with the dual-task approach has mostly used tasks that focused on 'learning that' (e.g. Gyselinck, Cornoldi, Dubois, de Beni, & Ehrlich, 2002; Kruley, Sciama, & Glenberg, 1994). Here, it has been shown that the visuo-spatial sketchpad is involved in processing text–picture combinations but not in processing text alone. These findings can be interpreted as suggesting that during processing text alone, no mental images were build, whereas during processing text and pictures, the pictures were processed in the visuo-spatial sketchpad. For example, Gyselinck et al. (2002) examined working memory processes using a spatial motor dual task as well as an articulatory dual task (which is assumed to load the phonological loop), while participants studied tasks that described basic notions of physics. The authors presented six tasks in text-only or multimedia format in a within-subjects design. The pictures represented the elements mentioned in the sentence, and the causal and temporal relationships between them. Learning outcomes tasks involved testing factual information that was explicitly given in the text and involved testing students' ability to draw inferences from several sentences. They found that performing the articulatory dual task interfered with learning from text-only and learning from text and pictures, indicating that text processing was hindered in both conditions by the articulatory dual task. On the other hand, the visuo-spatial dual task interfered with learning from text and pictures, however, not with learning from text only, implying that performing the visuo-spatial dual task interfered with picture processing, however, not with text processing.

A study investigating working memory involvement during procedural information processing was conducted by Brunyé, Taylor, Rapp, and Spiro (2006). Learners studied tasks that focused on how to assemble Kinder Egg™ toys in text-only, multimedia (i.e. text and picture) and picture-only formats. Learning outcomes tasks involved verifying whether two sequence steps appeared in the correct temporal order, describing each sequence and recalling in which format the procedure was studied. The results of Brunyé et al. (2006) also showed that the visuo-spatial dual task interfered with picture processing (in both the picture-only and multimedia format); however, it did not interfere with text processing. These findings seem to contradict our assumptions at first view, because the authors used tasks that at first sight seem to be 'learning how' tasks as described in the previous sections. However, importantly, in these tasks, no interactions between objects and body parts of the person performing the actions were described or depicted.[1] This implies that their tasks may not have afforded imagery.

To summarise, Gyselinck et al. (2002) and Brunyé et al. (2006) showed that performing a visuo-spatial dual task interfered with picture processing, however, not with text processing. These results imply that in 'learning that' tasks and 'learning how' tasks that do not seem to afford imagery, pictures are processed in the visuo-spatial sketchpad, whereas text is not. However, it is still unclear whether the involvement of working memory is the same when studying 'learning how' tasks that do afford mental imagery.

HYPOTHESES

The main purpose of this study was to investigate the involvement of the visuo-spatial sketchpad when studying 'learning how' tasks that afford mental imagery. On the basis of the assumptions and empirical results reported earlier, we made the following predictions.

First, in line with previous empirical findings (e.g. Mayer, 2009), we expected pictures to be beneficial for learning (i.e. a multimedia effect). Second, we expected the visuo-spatial sketchpad to be involved in learning in both instructional condition (as evidenced by interference of the dual task and learning), but for different reasons: In the text-only conditions, we expected the visuo-spatial sketchpad to be involved because procedural-motor tasks should trigger mental images (e.g. Fischer & Zwaan, 2008; Fogassi & Ferrari, 2011; Kosslyn et al., 2001; Postma & Barsalou, 2009; Rizzolatti & Craighero, 2004), which are assumed to be constructed in the visuo-spatial sketchpad (Baddeley, 1999). In the multimedia condition, we expected the visuo-spatial sketchpad to be involved because processing pictorial information loads the visuo-spatial sketchpad (e.g. Gyselinck et al., 2002; Kruley et al., 1994). Third, and most importantly, we expected an interaction between the dual task and the instructional format, indicating that the visuo-spatial sketchpad is less involved in processing text and pictures than in processing text alone. This can be assumed because constructing mental images from scratch should be more effortful than processing an external representation such as a picture (Larkin & Simon, 1987).

METHOD

Participants and design

Eighty-seven students from a Dutch university studied four first-aid procedures. Five participants were excluded, as dual-task performance showed that they had not followed the instructions properly. Of the remaining 82 participants (66 female and 16 male participants; $M = 22.80$ years, $SD = 3.05$), 78% had no prior experience with first-aid instructions. The experiment had a $2 \times 2 \times 2$ mixed design, with presentation format (i.e. text-only vs. multimedia) and dual task (i.e. with vs. without) as between-subject variables. We included time of testing (immediate vs. delayed) as a within-subject variable to see whether effects were the same for both immediate and delayed testing. Depending on presentation format and dual task condition, participants received (i) a text without dual task ($n = 21$), (ii) a text with a dual task ($n = 20$), (iii) a multimedia instruction without dual task ($n = 21$) or (iv) a multimedia instruction with a dual task ($n = 20$). Learning outcomes were measured immediately after learning and again after 1 week. Participants were randomly assigned to one of the four experimental conditions and received either payment or course credit for their participation.

Materials and measures

The procedural tasks were four first-aid instructions obtained from the Orange Cross manual (Henny, 2006). First-aid tasks were used as these tasks are procedural-motor tasks and pictures depicted the actor's body parts required to perform the actions. The tasks described (i) how to fold a sling in supporting a broken arm across a patient's chest, (ii) how to roll a patient from the recovery position onto their back, (iii) how to apply an easy-application bandage and (iv)

[1] Similar multimedia tasks that merely describe a procedure are, for example, operating a simulator (Booher, 1975), assembling a crane (Ellis, Whitehill, & Irick, 1996), operating an oscilloscope (Braby, Kincaid, Scott, & McDaniel, 1982), assembling a loading cart (Stone & Glock, 1981), operating a computer programme (Gellevij, van der Meij, de Jong, & Pieters, 2002) and setting a radio (Fukuoka, Kojima, & Spyridakis, 1999).

how to move an unconscious patient from areas of danger. The tasks contained 4, 5, 10 and 12 steps and 52, 51, 114 and 107 words, respectively. In the multimedia instructions, two or three pictures accompanied the text. In all tasks, the steps, which are required to perform the procedure correctly, and how these steps should be executed, were described in the text. Pictures were photographs in which both the object (i.e. the patient) and the person performing actions were depicted. In the text, action-related words were used (e.g. 'sit down in squat position with your feet at each side of the patient and as close to the patient as possible'). Studying the text alone allowed correct performance of the procedure.

The first-aid tasks were presented on a PC computer (Hewlett-Packard, Houston, USA) with 22 in. monitor. In the multimedia conditions, the text was presented to the left of the pictures, whereas in the text-only conditions, the text was presented in the middle of the monitor. The first-aid tasks were presented in random order using E-PRIME v. 1.2 (Schneider, Eschman, & Zuccolotto, 2002). The dual task used to load the visuo-spatial sketchpad during learning was the foot-tapping task (e.g. Liefooghe et al., 2005; Miyake et al., 2004; Service & Turpeinen, 2001). During the foot-tapping task, participants tapped four pedals on the floor with their right foot. They tapped one pedal per second in clockwise order. The size of the pedals was 6 cm wide and 9 cm long. They were arranged in a rectangle, which was 27 cm wide and 32 cm long.

To test how much knowledge participants had obtained, two retention tests were used. In the first test, which focused on procedural-motor/implicit knowledge, participants executed the first-aid tasks that they had studied using a first-aid dummy. This dummy was sitting in a chair for the two bandaging tasks and lying on the floor for the other two tasks. Bandage materials were provided. Participants did not receive any feedback on their performance. Performance accuracy was measured by the proportion of steps that were performed both correctly and in the correct order. Participants could score either 0 or 1 per step. The proportion of correctly performed steps was calculated, resulting in one score between 0 and 1. To assess the interrater reliability for performance accuracy, two raters coded 20% of all videos. Cohen's kappa was 0.71. The remaining 80% of the data were scored by a single rater only.

In the second test, which focused on declarative/explicit knowledge, participants saw a picture from the studied task and verbally described the steps that either preceded or followed the depicted step. Description accuracy was measured by the proportion of steps that were described both correctly and in the correct order. Participants could score either 0 or 1 per step. The proportion of correctly described steps was calculated, resulting in one score between 0 and 1. To assess the interrater reliability for description accuracy, two raters coded 20% of all texts. Cohen's kappa was 0.67. The remaining 80% of the data were scored by a single rater only.

As control variables dual task performance, learning time and prior knowledge were assessed. Dual task performance was gauged by assessing the randomness of the tapping sequence and measuring the time between taps. As a measure for randomness, the Phi-index was calculated using RGCalc (Towse & Neil, 1998). This index is a measure of repetition tendency and has a potential range between −100 (i.e. non-randomness) and 100 (i.e. randomness; Towse & Neil, 1998). As participants were required to tap pedals in a predefined sequence, a lower randomness score implies higher performance. Furthermore, the time between taps on the foot pedals was measured. As participants were required to tap one pedal each second, the deviance in milliseconds from 1 second could be calculated. Here, negative values indicate taps faster than 1 second, and positive values indicate taps slower than 1 second. Moreover, we assessed learning time to control for possible differences between groups regarding time dedicated to the learning tasks.

To assess the participants' prior knowledge concerning first-aid procedures, they had to indicate whether they had participated in a first-aid course (yes vs. no).

Procedure

Before the experiment started, participants gave their informed consent. Then, they familiarised themselves with performing a simple first-aid task that was unrelated to the experimental learning tasks on a first-aid dummy. Participants received a written task on paper, which described in three steps how to tilt a patient onto the side and back. Participants in the text-only conditions received this training task without a picture, whereas participants in the multimedia conditions received this task with a picture.

Subsequently, all participants answered a demographic questionnaire (i.e. age, sex, education and prior knowledge concerning first-aid procedures). Participants who had to perform a dual task were instructed to press the foot pedals in clockwise order and one pedal per second whilst studying the learning material. They were informed that if they would stop pressing the pedals for 5 seconds, a beep would remind them to continue pressing the pedals. After these instructions, they practised the foot-tapping task for 1 minute, while hearing a metronome that indicated the length of a second. Then, all participants were directed to study the four first-aid tasks and were informed that after studying these tasks, they would be tested on their acquired knowledge without referring back to the learning material. No time limit for studying was set so that participants could continue with the next task, by pressing a key on a keyboard, when they felt confident that they had understood the task and remembered all information. After learning the first-aid tasks, participants executed these tasks using the first-aid dummy. During their performance, participants' actions were recorded with a video camera from two angles. Also, a photograph was taken after the easy-application bandage had been applied around the first-aid dummy's arm. After executing a first-aid task, participants received a picture from the learning material and described the preceding or subsequent steps. This procedure was repeated for each task, in the same order as they had been studied. One week later, students performed the same learning outcome tests in the same order in the same room. No time limits were set for executing the post-tests. The first session took between 45 and 60 minutes, and the second session about 30 minutes. Each participant was tested individually.

Data analyses

To investigate whether performing a dual task interfered with processing 'learning how' tasks that afford mental imagery in a text-only or multimedia format, two mixed ANOVAs were performed. The dependent variables for these analyses were performance accuracy and description accuracy. The between-subject factors were presentation format (text-only vs. multimedia) and dual task (with vs. without). As within-subject factor, time of testing (immediate vs. delayed) was included. Partial eta-squared effect size is reported for interactions and main effects. For partial eta-squared, 0.01, 0.06 and 0.14 correspond to small, medium and large effect sizes, respectively (Cohen, 1988). For post-hoc pairwise comparisons, Cohen's d is reported as a measure of effect size. Here, 0.2, 0.5 and 0.8 correspond to small, medium and large effect sizes, respectively (Cohen, 1988). To follow up on significant interactions ($p < .05$), Bonferroni-adjusted pairwise comparisons were conducted.

RESULTS

In a first step, we tested whether the experimental conditions were similar with respect to the control variables assessed in the study. The participants of the four groups did not differ concerning age as revealed by a two-factorial ANOVA with presentation format (i.e. text-only vs. multimedia) and dual task (i.e. with vs. without) as independent variables, all Fs < 1. Furthermore, the four groups did not differ concerning gender, $\chi^2(3) = 0.89$, $p = .83$, or prior knowledge, $\chi^2(3) = 1.36$, $p = .72$.

A one-factorial ANOVA with presentation format as independent variable revealed that learners performing a dual task did not differ regarding randomness of the tapping sequence ($F < 1$; text-only condition: $M = -27.80$, $SD = 6.11$; multimedia condition, $M = -29.37$, $SD = 8.48$; note that all participants had negative values), nor for tapping time deviance ($F < 1$; text-only condition: $M = 220.12$, $SD = 98.96$; multimedia condition, $M = 210.18$, $SD = 134.19$), implying that participants in both conditions performed the dual task equally well.

Regarding learning time, a two-factorial ANOVA with presentation format and dual task as independent variables revealed no main effects of dual task ($F(1, 78) = 2.53$, $p = .12$, $\eta_p^2 = 0.03$; without dual task: $M = 109.36$, $SD = 44.45$; with dual task: $M = 96.45$, $SD = 26.57$) and presentation format ($F < 1$; text-only: $M = 105.03$, $SD = 44.12$; multimedia: $M = 101.10$, $SD = 29.06$). Furthermore, no interaction between presentation format and dual task was observed ($F(1, 78) = 2.18$, $p = .14$, $\eta_p^2 = 0.03$).

Main effects

Descriptive statistics of the dependent variables performance accuracy and description accuracy are provided in Table 1.

The results showed a main effect of presentation format for both performance accuracy and description accuracy, implying that participants in the multimedia condition performed and described the procedures correctly more often than participants in the text-only condition [performance accuracy: $F(1, 78) = 37.61$, $p < .001$, $\eta_p^2 = 0.33$, text-only condition: $M = 0.56$, $SD = 0.12$; multimedia condition, $M = 0.70$, $SD = 0.11$; description accuracy: $F(1, 78) = 5.42$, $p = .02$, $\eta_p^2 = 0.07$, text-only condition, $M = 0.33$, $SD = 0.18$; multimedia condition: $M = 0.41$, $SD = 0.18$]. These results imply that a multimedia effect was found for both dependent variables. However, the main effects of presentation format were qualified by interactions described subsequently.

The results also showed a main effect of dual task for both performance accuracy and description accuracy, implying that participants who did not perform the dual task performed and described the procedures correctly more often than participants who did perform the dual task [performance accuracy: $F(1, 78) = 21.93$, $p < .001$, $\eta_p^2 = 0.22$, without dual task: $M = 0.69$, $SD = 0.12$; with dual task, $M = 0.58$, $SD = 0.11$; description accuracy: $F(1, 78) = 13.97$, $p < .001$, $\eta_p^2 = 0.15$, without dual task: $M = 0.44$, $SD = 0.19$; with dual task, $M = 0.30$, $SD = 0.17$]. However, the main effect of dual task was qualified by interactions described subsequently.

The results did not show a main effect of time of testing for performance accuracy ($F(1, 78) = 2.65$, $p = .11$, $\eta_p^2 = 0.03$) but did show a main effect of time of testing for description accuracy, $F(1, 78) = 30.88$, $p < .001$, Cohen's $d = 0.49$, implying that description accuracy was larger for immediate ($M = 0.41$, $SD = 0.19$) than for delayed testing ($M = 0.32$, $SD = 0.18$).

Two-way interactions

The results showed a significant two-way interaction between presentation format and dual task for performance accuracy, $F(1, 78) = 9.84$, $p < .01$, $\eta_p^2 = 0.11$, however, not for description accuracy, $F(1, 78) = 2.37$, $p = .13$, $\eta_p^2 = 0.03$. Bonferroni-adjusted pairwise comparisons showed that performing a dual task interfered with information processing during learning with text alone ($p < .001$, $\eta_p^2 = 0.28$, Cohen's $d = 1.51$; text-only without dual task: $M = 0.65$, $SD = 0.13$; text-only with dual task: $M = 0.47$, $SD = 0.11$) but did not interfere with information processing during learning with multimedia ($p = .28$, $\eta_p^2 = 0.02$, Cohen's $d = 0.37$; multimedia without dual

Table 1. Means and standard deviations (in parentheses) for performance accuracy and description accuracy as a function of dual task, presentation format and time of testing

		Without dual task		With dual task	
		Text only	Multimedia	Text only	Multimedia
Performance accuracy in proportion correct	Immediately after learning	0.63 (0.13)	0.73 (0.10)	0.43 (0.10)	0.69 (0.12)
	One week later	0.67 (0.13)	0.70 (0.11)	0.51 (0.11)	0.67 (0.10)
Description accuracy in proportion correct	Immediately after learning	0.50 (0.23)	0.47 (0.17)	0.25 (0.16)	0.43 (0.18)
	One week later	0.35 (0.18)	0.43 (0.17)	0.20 (0.16)	0.31 (0.19)

task: $M = 0.72$, $SD = 0.11$; multimedia with dual task: $M = 0.68$, $SD = 0.11$).

The results also showed a significant two-way interaction between presentation format and time of testing for performance accuracy, $F(1, 78) = 14.97$, $p < .001$, $\eta_p^2 = 0.16$, but not for description accuracy, $F < 1$. Bonferroni-adjusted pairwise comparisons for performance accuracy showed that the multimedia effect was smaller for delayed testing than for immediate testing (immediate: $p < .001$, $\eta_p^2 = 0.40$, Cohen's $d = 1.58$; text-only: $M = 0.53$, $SD = 0.12$; multimedia: $M = 0.71$, $SD = 0.11$; delayed: $p < .001$, $\eta_p^2 = 0.16$, Cohen's $d = 0.88$; text-only: $M = 0.59$, $SD = 0.12$; multimedia: $M = 0.69$, $SD = 0.11$).

The results did not show a significant two-way interaction between dual task and time of testing [performance accuracy: $F(1, 78) = 1.62, p = .21, \eta_p^2 = 0.02$; description accuracy: $F < 1$].

Three-way interactions

The results showed a significant three-way interaction between presentation format, dual task and time of testing for description accuracy, but not for performance accuracy [performance accuracy: $F < 1$; description accuracy: $F(1, 78) = 7.27$, $p = .01$, $\eta_p^2 = 0.09$]. Bonferroni-adjusted pairwise comparisons for description accuracy showed that for immediate testing, performing a dual task affected learning outcomes only when learning with text alone (text-only: $p < .001$, Cohen's $d = 1.28$; without dual task: $M = 0.50$, $SD = 0.23$; with dual task: $M = 0.25$, $SD = 0.16$; multimedia: $p = .47$, Cohen's $d = 0.23$; without dual task: $M = 0.47$, $SD = 0.17$; with dual task: $M = 0.43$, $SD = 0.18$). However, for delayed testing, performing a dual task affected learning outcomes both when learning with text alone and when learning with multimedia (text-only: $p = .01$, Cohen's $d = 0.89$; without dual task: $M = 0.35$, $SD = 0.18$; with dual task: $M = 0.20$, $SD = 0.16$; multimedia: $p = .03$, Cohen's $d = 0.67$; without dual task: $M = 0.43$, $SD = 0.17$; with dual task: $M = 0.31$, $SD = 0.19$). When looking at this three-way interaction from a different point of view (Figure 1), the Bonferroni-adjusted pairwise comparisons showed that there

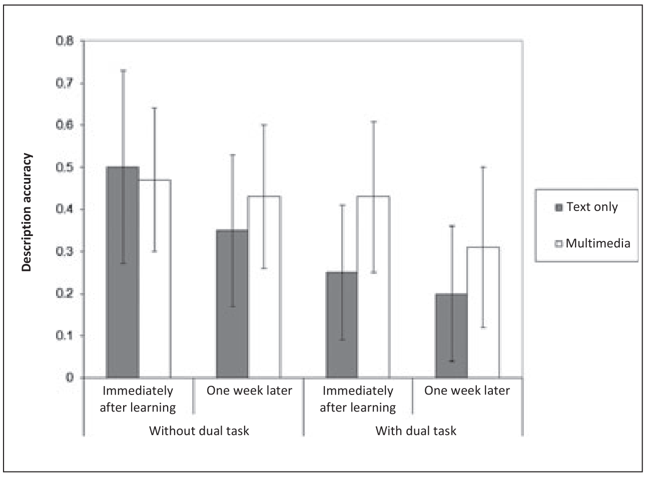

Figure 1. Means for the text-only and multimedia conditions for description accuracy as a function of time of testing and dual task (error bars represent standard deviations)

was a multimedia effect only for participants who learned with dual task immediate after learning (immediate/without dual task: $p = .71$, Cohen's $d = 0.15$; immediate/with dual task: $p < .01$, Cohen's $d = 1.07$; delayed/without dual task: $p = .15$, Cohen's $d = 0.46$; delayed/with dual task: $p = .06$, Cohen's $d = 0.63$).

DISCUSSION

The aim of the present study was to investigate whether the affordance for mental imagery in procedural-motor tasks (e.g. Fischer & Zwaan, 2008; Fogassi & Ferrari, 2011; Kosslyn et al., 2001; Postma & Barsalou, 2009; Rizzolatti & Craighero, 2004) has consequences for the processing of learning materials that consists of text only, or text and pictures in working memory. We focused on the role of the visuo-spatial sketchpad, because mental images and pictures are both assumed to be processed in this part of working memory (Baddeley, 1999).

The learning tasks used in this study described procedural-motor information concerning the position and movement of a person's body parts required to perform first-aid actions on a patient and were presented with text only, or with text and pictures. The dual task approach (i.e. learning with vs. without dual task) was used to assess the involvement of the visuo-spatial sketchpad during learning. Performance of the first-aid tasks was measured by performance accuracy, focusing on procedural-motor/implicit knowledge, and by description accuracy, focusing on declarative/explicit knowledge. Learning outcomes were measured immediately after learning and after 1 week. Additionally, dual-task performance and learning time were assessed to control for possible trade-offs between these variables and learning outcomes. No differences between groups were observed regarding dual-task performance and learning time, indicating that there was no trade-off between online and offline measures.

The results showed a multimedia effect for both learning outcome measures. However, for performance accuracy and for immediate testing of description accuracy, the results also showed that performance decreased when learning with text only because of performing the dual task, whereas performance was unaffected when learning with multimedia. Accordingly, the dual task interfered to a large extent with information processing during learning with text only but did not interfere with information processing during learning with multimedia. This implies that the visuo-spatial sketchpad was highly involved when learning with text only. These results are in contrast to previous research by Gyselinck et al. (2002), and Brunyé et al. (2006), who found that performing a visuo-spatial dual task interfered with picture processing whereas text processing did not.

A possible explanation for this discrepancy is that their tasks and our tasks differ in the degree to which the tasks afforded imagery. The tasks of Gyselinck et al. (2002) and Brunyé et al. (2006) probably did not afford imagery, as their tasks did not depict actions or did not depict actor's body parts required to perform the actions. In contrast, our 'learning how' tasks appear to have afforded mental imagery as they depicted the actor who performed actions on objects (i.e. patient). This could explain why in previous studies no evidence for the involvement of the visuo-spatial sketchpad during learning with text only was found. Taken together, this suggests that the role of the visuo-spatial sketchpad is different for learning with text only when tasks afford imagery compared with tasks that do not afford imagery.

In contrast to our assumptions, both for performance and immediate testing of description accuracy, the results showed that performing the dual task did not interfere at all with information processing when learning with multimedia. These results also differ from results obtained by Gyselinck et al. (2002), and Brunyé et al. (2006), who found that picture processing does require cognitive resources of the visuo-spatial sketchpad. A possible explanation for this discrepancy could be that the presentation of pictures in our study substituted the need for mental imagery and therefore strongly reduced the amount of cognitive resources of the visuo-spatial sketchpad required to understand the task.

Furthermore, the results for performance accuracy showed that the multimedia effect decreased over time. A similar pattern was found for description accuracy (multimedia effect with immediate testing, no multimedia effect with delayed testing) but only for participants who learned with dual task—no multimedia effect was found for participants learning without dual task. The effects of time of testing in this study have to be interpreted with care. First, participants were tested twice. According to the testing effect, retrieving information from memory positively influences memory for the task at a later time point (Carrier & Pashler, 1992) and therefore might also have affected performance after 1 week. Second, participants who learned with text only saw a picture from the learning material during the description post-test. As participants who learned with multimedia already saw the picture during the learning phase, this picture may have positively influenced memory for the task only for participants who learned with text only and therefore may have affected their performance after 1 week. If this explanation is correct, this also shows that pictures are beneficial to learning. In this study, it is impossible to assess how these issues affected the pattern of results in each experimental condition.

Finally, the results for performance accuracy and description accuracy do not show exactly the same pattern of results. For example, the results for description accuracy did not show a multimedia effect for participants who learned without dual task. A possible explanation for this missing multimedia effect is that, when 'learning how', pictures are suitable to convey procedural-motor/implicit knowledge, as gauged by performance accuracy, however, are less suitable to convey declarative/explicit knowledge, as gauged by description accuracy.

Limitations and future research

One limitation of the conducted study lies in the fact that we used only tasks describing 'learning how' but no tasks describing 'learning that'. Using both kinds of tasks in one experiment (i.e. as independent variable) can provide much stronger evidence that the observed interference between the spatial motor dual task and processing text only is unique to tasks with high imagery affordance (i.e. procedural-motor tasks). Therefore, future studies should include causal as well as procedural tasks, which should ideally only differ regarding their affordances for imagery but are nonetheless comparable regarding other variables, which might also influence learning outcomes (e.g. text difficulty, text–picture correspondence and number of propositions).

Another drawback of the conducted study lies in the fact that we did not use another dual task to be able to differentiate between the general impacts of performing a dual task on learning outcomes versus the specific impact of performing a dual task that taps on the respective subsystems of working memory. One might argue that it is possible to make the *a priori* assumption that spatial motor dual tasks interfere with information processing in the visuo-spatial sketchpad, because there is a large corpus of evidence in the context of working memory research, indicating that spatial motor dual tasks interfere with information processing in the visuo-spatial sketchpad and do not interfere with information processing in other working memory subsystems (e.g. Farmer et al., 1986; Postle, Idzikowski, Della Sala, Logie, & Baddeley, 2006; Quinn & Ralston, 1986). Nonetheless, we think that using another dual task would strengthen our results, especially because we used procedural learning tasks in our study, which—to our knowledge—have never been used in combinations with the secondary task paradigm before.

A third drawback of the reported study lies in the fact that we conducted only one single experiment. Thus, the current paper is only a first contribution to the question whether the affordance for mental imagery in procedural-motor tasks has consequences for the processing of learning materials that consist of text alone, or text and pictures in working memory.

Moreover, from a theoretical point of view, the question whether pictures are processed at 'zero costs' (Larkin & Simon, 1987, p. 92) cannot be answered with complete confidence from this experiment. Therefore, it is essential that future research investigates the involvement of the visuo-spatial sketchpad in multimedia learning tasks focusing on 'learning how' that afford imagery to different extents or by instruction participants to apply imagery.

Overall, the results of the reported experiment provide first evidence that pictures are beneficial when 'learning how', because they reduce the cognitive effort that is required to understand the procedural learning material. By reducing the required cognitive effort for processing one part of the learning task, more effort can be spent on other parts of the learning task. Even though Carney and Levin (2002) argued that presenting pictures is superfluous when text elicits mental images in students, the results from this study suggest that using pictures when a task affords imagery can facilitate understanding of the learning material.

To conclude, pictures seems to be beneficial to learning, as pictures in tasks that focus on 'learning how' omit the need to engage in imagery and therewith reduce the cognitive effort that is required to understand the learning material. However, it should be noted that this is especially the case when tasks afford imagery by describing the position and movement of an actor's body parts required to perform actions.

ACKNOWLEDGEMENTS

This research is funded by the Pact for Research and Innovation of the Competition Fund of the Leibniz Gemeinschaft and by the Center of Advanced Media Studies Amsterdam. We would like to thank the Dutch Orange Cross and the German Red Cross for providing instructional materials. We would also like to thank Katrin Schleinschok, Marion Kornmayer and Laura Seyfried for their assistance in scoring the data.

REFERENCES

Andrade, J. (2001). *Working memory in perspective*. Hove: Psychology Press.

Baddeley, A. D. (1999). *Essentials of human memory*. Hove: Psychology Press.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. DOI: 10.1146/annurev-psych-120710-100422

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*, 126–145. DOI: 10.1037/0096-3445.129.1.126

Baddeley, A. D., Grant, S., Wight, E., & Thomson, N. (1973). Imagery and visual working memory. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V* (pp. 205–217). London: Academic Press.

Booher, H. R. (1975). Relative comprehensibility of pictorial information and printed words in proceduralized instructions. *Human Factors*, *17*, 266–277.

Braby, R., Kincaid, J. P., Scott, P., & McDaniel, W. C. (1982). Illustrated formats to teach procedures. *IEEE Transactions on Professional Communication*, *25*, 61–66. DOI: 10.1109/TPC.1982.6447756

Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology*, *20*, 917–940. DOI: 10.1002/acp.1236

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*, 5–26. DOI: 10.1023/A:1013176309260

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, *20*, 633–642. DOI: 10.3758/BF03202713

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

Dechsri, P., Jones, L. L., & Heikkinen, H. W. (1997). Effect of a laboratory manual design incorporating visual information-processing aids on student learning and attitudes. *Journal of Research on Science and Teaching*, *34*, 891–904. DOI: 10.1002/(SICI)1098-2736(199711)34:9<891::AID-TEA4>3.0.CO;2-P

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool of unwelding visuo-spatial memory. *Neuropsychologia*, *37*, 1189–1199. DOI: 10.1016/S0028-3932(98)00159-6

Ellis, J. A., Whitehill, B. V., & Irick, C. (1996). The effects of explanations and pictures on learning, retention, and transfer of a procedural assembly task. *Contemporary Educational Psychology*, *21*, 129–148. DOI: 10.1006/ceps.1996.0012

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *Quarterly Journal of Experimental Psychology*, *38*, 675–688. DOI: 10.1080/14640748608401620

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, *61*, 825–850. DOI: 10.1080/17470210701623605

Fogassi, L., & Ferrari, P. F. (2011). Mirror systems. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 22–38. DOI: 10.1002/wcs.89

Fukuoka, W., Kojima, Y., & Spyridakis, J. H. (1999). Illustrations in user manuals: Preference and effectiveness with Japanese and American readers. *Technical Communication*, *46*(2), 167–176.

Gellevij, M., van der Meij, H., de Jong, T., & Pieters, J. (2002). Multimodal versus unimodal instruction in a complex learning context. *The Journal of Experimental Education*, *70*, 215–239. DOI: 10.1080/00220970209599507

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

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, *31*, 129–151. DOI: 10.1016/0749-596X(92)90008-L

Gyselinck, V., Cornoldi, C., Dubois, V., de Beni, R., & Ehrlich, M.-F. (2002). Visuospatial memory and phonological loop in learning from multimedia. *Applied Cognitive Psychology*, *16*, 665–685. DOI: 10.1002/acp.823

Hayes, D. A., & Henk, W. A. (1986). Understanding and remembering complex prose augmented by analogic and pictorial illustration. *Journal of Literacy Research*, *18*, 63–78. DOI: 10.1080/10862968609547556

Henny, W. (2006). *Orange Cross Manual, official manual for providing first aid* (25th edn). Utrecht: ThiemeMeulenhoff.

Iserbyt, P., Mols, L., Elen, J., & Behets, D. (2012). Multimedia design principles in the psychomotor domain: The effect of multimedia and spatial contiguity on students' learning of basic life support with task cards. *Journal of Educational Multimedia and Hypermedia*, *21*(2), 111–125.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*, 635–642. DOI: 10.1038/35090055

Kruley, P., Sciama, S. C., & Glenberg, A. M. (1994). On-line processing of textual illustrations in the visuospatial sketchpad: Evidence from dual-task studies. *Memory & Cognition*, *22*, 261–272. DOI: 10.3758/BF03200853

Langston, W., Kramer, D. C., & Glenberg, A. M. (1998). The representation of space in mental models derived from text. *Memory and Cognition*, *26*, 247–262. DOI: 10.3758/BF03201137

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–99. DOI: 10.1111/j.1551-6708.1987.tb00863.x

Lawrence, B. M., Myerson, J., Oonk, H. M., & Abrams, R. A. (2001). The effects of eye and limb movements on working memory. *Memory*, *9*, 433–444. DOI: 10.1080/09658210143000047

Liefooghe, B., Vandierendonck, A., Muyllaert, I., Verbruggen, F., & Vanneste, W. (2005). The phonological loop in task alternation and task repetition. *Memory*, *13*, 550–560. DOI: 10.1080/09658210444000250

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: visual, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 105–115). Amsterdam: Elsevier.

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*, 2172–2178. DOI: 10.1016/j.chb.2013.04.035

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*, 49–63. DOI: 10.1037/0022-0663.88.1.49

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.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*, 715–726. DOI: 10.1037//0022-0663.82.4.715

Michas, I. C., & Berry, D. C. (2000). Learning a procedural task: Effectiveness of multimedia presentations. *Applied Cognitive Psychology*, *14*, 555–575. DOI: 10.1002/1099-0720(200011/12)14:6<555::AID-ACP677>3.0.CO;2-4

Millar, R., Lubben, F., Gott, R., & Duggan, S. (1994). Investigating in the school science laboratory: Conceptual and procedural knowledge and their influence on performance. *Research Papers in Education*, *9*, 207–248. DOI: 10.1080/0267152940090205

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, *115*, 123–142. DOI: 10.1016/j.actpsy.2003.12.004

Postle, B. R., Idzikowski, C., Della Sala, S., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 100–120. DOI: 10.1080/17470210500151410

Postma, A., & Barsalou, L. W. (2009). Spatial working memory and imagery: From eye movements to grounded cognition. *Acta Psychologica*, *132*, 103–105. DOI: 10.1016/j.actpsy.2009.07.006

Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *Quarterly Journal of Experimental Psychology*, *38*, 689–703. DOI: 10.1080/14640748608401621

Rizzolatti, G., & Craighero, L. (2004). The mirror–neuron system. *Annual Review of Neuroscience*, *27*, 169–192. DOI: 10.1146/annurev.neuro.27.070203.144230

Ryle, G. (2000). *The concept of mind*. Chicago: University of Chicago Press.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.

Schüler, A., Scheiter, K., & van Genuchten, E. (2011). The role of working memory in multimedia instruction: Is working memory working during learning from text and pictures? *Educational Psychology Review*, *23*, 389–411. DOI: 10.1007/s10648-011-9168-5

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction*, *14*, 293–305. DOI: 10.1016/j.learninstruc.2004.06.005

Service, E., & Turpeinen, R. (2001). Working memory in spelling: Evidence from backward typing. *Memory*, *9*, 395–421. DOI: 10.1080/09658210143000137

Smyth, M. M., & Pendleton, L. R. (1990). Space and movement in working memory. *The Quarterly Journal of Experimental Psychology Section A*, *42*, 291–304. DOI: 10.1080/14640749008401223

Stone, D. E., & Glock, M. D. (1981). How do young adults read directions with and without pictures? *Journal of Educational Psychology*, *73*, 419–426. DOI: 10.1037/0022-0663.73.3.419

Towse, J. N., & Neil, D. (1998). Analyzing human random generation behavior: A review of methods used and a computer program for describing performance. *Behavior Research Methods, Instruments, & Computers*, *30*, 583–591. DOI: 10.3758/BF03209475

Van Genuchten, E., Scheiter, K., & Schüler, A. (2012). Examining learning from text and pictures for different task types: Does the multimedia effect differ for conceptual, causal, and procedural tasks? *Computers in Human Behavior*, *28*, 2209–2218. DOI: 10.1016/j.chb.2012.06.028

Van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror–neuron system and observational learning: Implications for the effectiveness of dynamic visualizations. *Educational Psychology Review*, *21*, 21–30. DOI: 10.1007/s10648-008-9094-3

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

Wan, F. M., & Baragash, R. S. A. (2011). Dynamic visualization on student's performance, retention and transfer of procedural learning. *World Academy of Science, Engineering and Technology*, *59*, 1249–1253.