Information modalities for procedural instructions

The influence of text, pictures and film clips on learning and executing RSI exercises

Charlotte van Hooijdonk & Emiel Krahmer

Communication and Cognition

Faculty of Arts

Tilburg University

The Netherlands

Abstract

Much of the empirical research on the effectiveness of different instructional designs has focussed on declarative tasks, where a learner acquires knowledge about a certain topic. It is unclear to what extent findings for learning declarative tasks (which are not consistent on all aspects), carry over to learning procedural tasks, where a learner acquires a certain skill. Moreover, it may well be that besides variation between domains, variation within domains also influences the relative contribution of the instruction format to the learning of a task. In this paper, we describe two experiments studying a specific kind of procedural instructions, namely exercises for the prevention or Repetitive Strain Injury (RSI), taking information modality (text vs. picture vs. film clip) and difficulty degree of the exercises (easy vs. difficult) into account. In the first experiment, participants had to learn RSI exercises and were asked to execute them. The results showed that both information modality and difficulty degree had an effect on learning times, number of practiced exercises during learning, execution times, and number of correctly executed exercises. No differences were found in the subjective satisfaction of the participants regarding the presentation of the exercises. A second (forced choice) experiment was conducted to determine participant's preference for one of these three instructional formats. The results showed that most participants preferred learning from film clips rather than from text or pictures.

Introduction

The emergence of computer-based learning has led to new possibilities for presenting instructions and to a renewed interest in the effects of different information modalities. Instructions can be given in plain text, but also in the form of static visuals (e.g., diagrams, pictures) or in dynamic visuals (e.g., animations, film clips). Naturally, this raises the question which (combinations of) modalities are best for which learning task. This question has been addressed in a large number of recent studies (e.g., Lewalter 2003, Lowe 2004, Mayer 2001, 2003, Michas & Berry 2000, Ploetzner & Lowe 2004, among many others).

Each of the aforementioned presentation modes has its own unique characteristics, and its own advantages and disadvantages from an instructional perspective. Language is the basic form of human communication, and one of its main strengths is that it is expressive and explicit (both for concrete and for abstract subject matter). An additional advantage of its written form (as opposed to the spoken variant) is that it is not transient: written sentences remain visible on paper and can be re-read if desired, while spoken sentences are 'gone' once they have been uttered. But, reading a text requires considerable skill and effort, and, moreover, text primarily facilitates linear information processing. In a seminal paper, Larkin & Simon (1987) compare textual representations with pictures offering diagrammatic representations, and argue that the latter are beneficial for learning because they visualize locations and make information explicit that is often left implicit in textual representations (e.g., relations between locations). They claim that diagrams are 'computationally' more efficient than texts, in that they support human search and recognition, both factors that contribute to learning, and this claim has found considerable empirical support (see e.g., Tversky et al. 2002 for a review).

The focus of many recent studies has been on the instruction effects of *dynamic* visuals, presumably because such instructions only recently have become a real possibility due to the increased computing power of standard multimedia PCs. A number of reasons have been suggested why an advantage of dynamic visuals over other presentation possibilities such as text and static visuals might be expected. For instance, it has been argued that dynamic visuals are beneficial for learning since they offer a “complete model” of a learning task (e.g., Park & Hopkins, 1993; Lewalter 2003); in other words, they are “informationally complete” (Schnotz et al. 1999:249) When static visuals or text is used, learners themselves will have to construct a mental representation of the temporal aspects in the instruction. It has been argued that dynamic illustrations offer a better representation of these temporal aspects, thereby reducing the level of abstraction, and supporting a deeper level of understanding than static visuals would do (Park & Hopkins, 1993). Arguably, this should facilitate learning, since it would reduce learning times, would require less practice, and would result in fewer errors.

A substantial number of studies have tried to demonstrate this presumed learning benefit, however with mixed results (e.g., Betrancourt & Tversky 2000, Lewalter 2003, Michas & Berry 2000, Tversky et al. 2002). Two groups of explanations for these mixed results have been offered: one group of explanations stresses the disadvantages of dynamic visuals, another group points to methodological problems in these studies. An explanation of the first kind would be that dynamic visuals have a fixed duration, which viewers simply have to watch. This may lead to inherently longer learning times (Tversky et al. 2002). In a similar vein, it has been suggested that dynamic visuals may take more time to *process* than other presentation modes; the information in dynamic visuals changes continuously, and as a result learners could be overwhelmed with the amount of information (Lewalter, 2003). Alternatively, it has been suggested that dynamic visuals are processed somewhat superficially; they require little cognitive effort, as they can be watched rather passively (e.g., Schnotz et al. 1999, Schnotz & Rasch 2005). Explanations of the second kind, by contrast, argue that the mixed findings so far are primarily due to methodological problems in the various experimental studies. For instance, to be really beneficial dynamics should have some added value (e.g., Weiss et al. 2002), which is not always the case. When it comes to temporal sequencing or indicating direction of motion, it has been claimed that arrows in static visuals may be just as effective as dynamic visuals (e.g., Tversky et al. 2000). Tversky and co-authors point out that in some comparative studies there is a lack of equivalence between dynamic and static visuals in content or procedures, for instance because the dynamic visuals convey more information or involve interactivity. Some researchers have argued that even when dynamic visuals do not lead to improved learning, they are more attractive and motivating than other instruction forms, and should be preferred for that reason (e.g., Perez & White 1985, Rieber 1991, see also Tversky et al. 2002). However, this "subjective satisfaction" (Nielsen 1993) of various instruction modes is often not addressed, and when it is the results are equivocal (e.g., Pane et al. 1996).

A further complicating factor is that the learning domain itself might also have an influence on the relative benefits of information modalities (Weiss, Knowlton & Morrison, 2002). It seems reasonable to assume that different learning domains benefit from different kinds of instructions (cf., also Hegarty 2004). In many studies, dynamic visuals are used as learning instructions for descriptive, "scientific" explanations (often describing causes and effects), for instance, of mechanical and electronic systems (e.g., brakes, pumps, generators, Mayer 1989, Mayer and Gallini 1990; electronic circuit systems, Park and Gittelman, 1995), mathematical concepts (e.g., algebra, Reed 1985), or complex natural phenomena (e.g., electricity, Cheng 2002; lightning, Mayer and Moreno 2002; gravitational lensing, Lewalter 2003; meteorological changes, Lowe 2004). Arguably, some of these scientific learning domains lend themselves better for dynamic visualization than others. Moreover, what holds for descriptive tasks (such as those above), may not apply to procedural ones, such as bandaging a hand (Michas & Berry 2000), folding origami models (Caroll & Wiebe 2004) or tying nautical knots (Schwan & Riemp 2004). These procedural tasks differ from descriptive ones in a number of respects. Not only is the nature of the task different (procedures are inherently more linear, one step following another), but also the learning goal is different (the focus is not on understanding, but on acquiring certain capabilities or skills). To further complicate the situation it may well be that besides variation *between* domains there is also variation *within* domains. Arguably, some learning tasks in a given domain are easier than others, and this may have an influence on the relative contribution of various instruction formats for those tasks.

In this paper we report on a study of the effects of task difficulty and information modality (comparing dynamic visuals with static visuals and text) on learning a specific class of procedural tasks, namely exercises aiming at the prevention of Repetitive Strain Injury (RSI). RSI is a general term for disorders that are caused by repetitious use of hands, arms, and shoulders, often as a result of prolonged computer terminal work (e.g., Stone 1983). Recently, the term CANS (short for Complaints of Arm, Neck and/or Shoulders) was introduced as an alternative term, but here we will stick to the more familiar RSI terminology. It is generally assumed that taking regular breaks during computer work in combination with exercises is beneficial for the prevention of RSI (e.g., Balci & Aghazadeh 2003, McLean et al. 2001, Williams et al. 1989). RSI exercises offer a new and understudied learning domain, with various interesting properties. Generally, these exercises involve little or no abstraction, do not consist of a complicated sequence of actions, and are relatively short. Moreover, the exercises are highly recognizable (almost everybody has two hands). It is interesting to observe that current RSI information websites offer a large variety of prevention exercises, in many different presentation formats (see Figure 1 for a representative selection) which raises the natural question what the effectiveness of the various presentation formats is. Note also that there is substantial variation in the difficulty level of existing RSI exercises, so that the other factor of interest (variation in task difficulty) can be modelled in a fairly straightforward way in this domain.

[INSERT FIGURE 1]

The potential influence of both presentation modality and task difficulty on learning performance can be motivated from *cognitive load theory* (e.g., Sweller & Chandler 1991, Marcus et al. 1996, Sweller et al. 1998, Sweller 1999), a theory which aims to develop optimal instructional designs while considering the limitations of the human mind. Cognitive load theory builds on the broadly accepted assumption that the mind combines a short term (or working) memory of very limited capacity (where all conscious activity and processing of information occurs; Miller 1956, Baddely 1992) and a long term memory with a virtually unlimited capacity (e.g., Sweller et al. 1998). According to Cognitive load theory, learning amounts to the construction of new (or modification of existing) *schemata* (Chi et al. 1982), which are subsequently stored in long term memory. Since the capacity of working memory is severely limited, the cognitive load of learners should be kept at a minimum during learning. In the current version of the theory (Sweller et al. 1998, Sweller 1999), three kinds of load are distinguished: *intrinsic* load, caused by the structure and intrinsic nature of the learning task, *extraneous* load, caused by the manner of presentation and its influence on working memory, and *germane* load, caused by learners' effort to process and comprehend learning material. The sum of these three kinds of load should not exceed working memory capacity, in order to avoid cognitive overload. As argued above, the intrinsic load may vary both between and within task domains. Since the intrinsic load of a particular learning task is fixed, instruction design can only influence the extraneous and the germane load, and, obviously, when the intrinsic load of a particular task is high, there is relatively less room for extraneous and germane load.

Germane load is a 'positive' form of load, since it encourages learners to engage in cognitive processing that may lead to improved schemata construction (Sweller et al. 1998, van Merriënboer et al. 2002). This suggests that optimal instructions are those which minimize extraneous and maximize germane load, while simultaneously avoiding overload. A complicating factor is that the distinction between extraneous and germane cognitive load, although intuitively plausible, cannot reliably be measured (van Merriënboer et al. 2002). In general, measuring cognitive load in realistic learning situations is not straightforward. While some recent attempts have been made to measure cognitive load directly (e.g., the dual task approach advocated by Brünken et al. 2003), many studies rely on indirect measures such as behavioral or learning outcome measures (see e.g., Brünken et al. 2003, van Merriënboer et al. 2002). Although these measures only relate to cognitive load indirectly, they do assess learning behavior directly which is sufficient for our current purposes. In this study, we concentrate on learning times, amount of practising during learning, execution times, and execution errors.

Arguably, the different information modalities of interest (text, picture, film clip) have different loading potentials, which may influence performance on one or more of these measures. Of all three modalities, it seems likely that text imposes the highest load: it requires more mental effort to read a text than to watch a picture or a film clip, hence it seems likely that learning from text takes longer than learning from pictures or film clips. Potentially, an additional complication for learning RSI prevention exercises from text is that schemata construction may be more involved than for pictures and film clips. In the text case, learners have to 'translate' the textual instructions to manual gestures. Notice that this is a concrete instance of Glenberg's Indexical Hypothesis, stating that readers associate words and phrases with objects and actions in 'the real world', which should facilitate understanding (e.g., Glenberg 1997, Glenberg and Robertson 1999). An instruction in the form of a film clip and (probably to a lesser extent) a picture, depicts the hand and arm movements the learner should make, while the learner has to infer these gestures when presented in text. Hence, it is hypothesized that learners will practice more often while learning from text than while learning from visual presentations. An interesting question is whether the load imposed by text is only extraneous or also partly germane. It might be that the extra effort required to learn from text leads to good learning results (short execution times, few execution errors), especially when the intrinsic load is low (so that overload can be avoided).

Pictures arguably impose the lowest load of the three presentation modalities; viewing a static picture requires little mental effort. Since a picture captures the essential information of a procedure, it is to be expected that learning times for pictures are relatively short. However, due to their static nature, pictures offer little information about the temporal structure of a procedure, and for more complicated exercises (i.e., exercises with a higher intrinsic load) this may hamper schemata construction and might result in suboptimal execution performance.

To what extent film clips impose cognitive load is uncertain: on the one hand, it can be argued that they may induce load, since the film clips continuously change and learners have to process this information which reduces the cognitive resources available for germane load, but they may also lessen cognitive load by relieving the learner in the translation process, which is required when reading text. The main strength of film clips might well be their "informational completeness"; learners do not have to infer the exact sequence of movements from text or from a single snapshot, the entire sequence of actions is visualized, which arguably facilitates schemata construction. This suggests that learners will not practice much during learning. Whether this will also result in few execution errors is not obvious: it may be that the large amount of external support relieves learners of cognitive load demands that they would be able to fulfil, but which might prevent them from performing beneficial cognitive actions for learning.

In this paper we describe two experiments. In the first experiment (section 2), participants learn and execute 20 RSI prevention exercises in two degrees of difficulty. We measure the influence of presenting an instruction in text, picture or film clip on learning times, amount of practicing during learning, execution times, and number of execution errors. Besides these objective measures, participants are also asked for their subjective satisfaction. In the second experiment (section 3), participants are asked which instructional format they preferred in a forced choice experiment. This experiment basically tries to find learning preferences, in a manner somewhat similar to Leutner & Plass (1998), on the visualizer-verbalizer dimension (Kirby & Moore 1988). We end with a general discussion in section 4.

Experiment I: efficiency, effectiveness, and subjective satisfaction of information modalities

Method

Participants

Participants were 30 young adults, between 18 and 30 years of age. Of the participants, 15 were male and 15 were female.

Design

The experiment had a 3 (information modality) × 2 (difficulty degree) factorial design, with information modality (dynamic visual [film clip], static visual [picture], text) as between participants variable and difficulty degree (easy, difficult) as within participants variable, and with learning times, amount of practicing during learning, execution times, and number of correctly executed exercises as dependent variables. The participants were randomly assigned to an experimental condition.

Stimuli

A total of 20 RSI prevention exercises were chosen from web sites on RSI prevention and RSI injury prevention software[1]. The chosen exercises were all

[1] http://web.mit.edu/atic/www/disabilities/rsi/exercises.html
http://www.workpace.com

exercises to prevent RSI in hands and arms. Of the 20 exercises, ten exercises were assumed to be easy to perform, and ten were assumed to be difficult. The criterion for a difficult exercise was that it should be either a complex symmetrical movement or an asymmetrical movement. Complex symmetrical movements were classified as movements consisting of at least three sequential atomic movements, in the course of which both arms and hands make the same movements. Asymmetrical movements were classified as movements in which the participant should make a different movement with each arm or hand. Easy exercises were neither asymmetrical nor complex movements. Figure 2 shows representative examples of an easy and a difficult exercise.

[INSERT FIGURE 2]

Note that this operationalization of easy and difficult tasks is based on the relative complexity of the sequence of motoric movements. To find out to what extent this objective criterion coincided with subjective perception of task difficulty a pre-test was carried out, in which 9 participants were asked to classify the 20 exercises (presented in text+picture format, and in random order). They were instructed to make two piles, the first consisting of the ten exercises they considered easiest to perform, the second pile consisting of the ten exercises they considered more difficult to perform. It turned out that of the 10 exercises classified as easy, 7 were indeed perceived as easy by the vast majority (> 75%) of the participants, while the remaining 3 were perceived as difficult by a majority of the participants (presumably because these consisted of gestures that are motorically simple but not commonly used and hence with which participants may have been less familiar). Of the 10 exercises classified as difficult, 9 were indeed perceived as such by the vast majority (> 75%), while the remaining one was perceived as easy by most participants (interestingly, this was an exercise that was motorically complex but familiar to most participants). In sum, for 16 of the 20 exercises there was a clear and consistent correlation between the objective and the perceived difficulty degree. Throughout this paper, we will report results relating to the original objective classification of movements (also see discussion).

The 20 RSI prevention exercises were presented in three different formats. In the text condition, the exercises were explained to the participants in a purely textual formal. The total amount of words did not differ between the 10 easy and 10 difficult exercises: both counted 268 words in total. Thus, the average length was almost 27 words per exercise. Since some exercises are a few words shorter and others a few words longer, we will only report on mean *total* averages for the 10 exercises in each condition. Because natural language is often ambiguous, the text exercises were checked on their comprehensibility in a second pre-test with three participants (different from those in the first pre-test). It turned out that a few exercises lead to misunderstandings and these exercises were reformulated. The new versions were presented to the same three participants, and they agreed that the reformulations solved the misunderstandings.

In the picture condition, each of the 20 exercises was displayed in a single photograph. For this, pictures were taken with a digital camera from a female who made the exercises. She wore black clothing and the movements were shot against a black background so that only her hands were visible on the picture. The photo depicted the *stroke* of the movement, which is that phase of the movement as it unfolds in time containing the ‘semantic content’ of the movement (Kendon, 1980). To indicate the direction of movement, arrows were added to the pictures of nine difficult and two easy exercises. The size of the pictures was 1536 * 1014 pixels.

For the film clip condition, the same female in an identical surrounding was filmed with a digital film camera (25 frames per second). The total amount of frames did not differ between easy and difficult exercises: both counted 1097 frames (average film length was thus 4.39 seconds). Again, since some film clips are somewhat shorter, and others somewhat longer, we will report on mean total averages for the 10 easy and the 10 difficult exercises.

The exercises were presented to the participants via a web site: one web site for each condition. The experiment was run on a multimedia PC, with a 17-inch color monitor. In all three conditions, the presentations of the RSI exercises appeared at the centre of the display. In the Video web site the movements were shown in a film clip with a start and a stop button and a slider. The participants had the possibility to view the film as often as they desired, but not much use was made of this option. Exercises were presented in one of two random orders to control for possible learning effects.

Procedure

The experiment included three parts: a practice session, the central part in which the 20 RSI exercises were presented, and finally, a questionnaire. The participants took part one at a time. Each participant was invited to an experimental laboratory, and took a seat behind the computer. Participants were told that they would receive 20 exercises which they had to learn and perform to assess to what extent they suffered from RSI. In addition, they were led to believe that their hand and arm movements were filmed because a doctor would later look at the recordings of the participants to determine to what extent they suffered from RSI. The procedure of the experiment is depicted in figure 3.

[INSERT FIGURE 3]

After the participants had read the instructions, they could click on the hyperlink "start", and the first trial exercise appeared. Depending on their experimental condition, the participants read or viewed the trial exercise until they thought that they could properly execute the exercise. Subsequently they clicked the hyperlink 'next' at the bottom of the page. A new webpage appeared with the text "execute trial exercise 1" at the centre of the page. During the execution of an exercise, participants could not see the instruction. When they had executed the exercise, they clicked the hyperlink "next exercise" at the bottom of the page. After completing the second trial exercise, participants were asked if they had any questions regarding the experimental procedure. If not, the participant could start the actual experiment, proceeding in the same way as during the trial session. During the experiment, there was no further interaction between participants and experiment leader.

After the participants completed the 20 experimental exercises, they received a questionnaire. In this questionnaire the subjective satisfaction regarding the content and structure of the web site as well as the comprehensibility and the attractiveness of the exercises were measured. The questionnaire consisted of 16 bi-polar 7-point semantic differentials addressing structure and content of the websites as well as comprehensibility and attractiveness of the exercises (see appendix A). The presentation order of the adjectives was randomized. For processing the positive adjectives were mapped to 1 (= very positive) and the negative ones to 7 (= very negative). The participants were debriefed at the end of the experiment.

Data processing

The following data were collected: learning times, number of practiced exercises during the learning time, execution times, number of correctly executed exercises, and the results of the questionnaire.

The time it took the participants to learn and execute the exercises was computed from the data of the log program ProxyPlus[2]. This program registered the times associated with participants' mouse clicks during the experiment. The time period between clicking the hyperlink "next" which preceded the instruction of an exercise and the hyperlink "next" which followed the instruction of an exercise was defined as the *learning time* (see figure 3). The time period between clicking the hyperlink "next" which followed the instruction of an exercise and the hyperlink "next" that preceded a new instruction of an exercise was defined as the *execution time* (see figure 3).

A digital video camera was used to record the movements the participant made during the experiment. These video recordings of the participants' hands and arms were used to assess whether the participants practised the exercise during the learning period and to assess their performance while executing the RSI exercises. Occasionally, a participant performed an exercise in a correct but not intended way. These cases were counted as correctly executed exercises. One judge did the assessment: scoring was easy, and the few difficult cases that did arise were resolved after discussion.

Tests for significance were performed using a repeated measures analysis of variance (ANOVA), with a significance threshold of .05. For post hoc tests the Tukey

[2] http://www.proxyplus.com/

HSD method was used. The internal consistency of the four item sets of the questionnaire was measured using Cronbach's α.

Results

[INSERT TABLE 1]

Learning times

Table 1 summarizes the results. First consider the learning time. It was found that the difficulty degree had an effect on the amount of time to learn the exercises ($F (1,27) = 37.35$, $p < .001$). Overall, the participants needed more time to learn the difficult exercises than the easy ones. Also the information presentation had an effect on the learning time ($F (2,27) = 4.53$, $p < .025$). Participants in the picture condition required the shortest learning times, participants in the text condition had the longest learning times, with learning times from film clips in between these two. Post-hoc tests indicated that the instruction in text differed significantly from the instruction in a picture ($p < .025$). There was no significant difference between the instruction in text and film clip ($p = .35$). Also, the instruction in a picture did not differ significantly from the instruction in a film clip ($p = .25$). No significant interactions between difficulty degree and information modality were found.

Amount of practising during the learning period

Table 1 also reveals that the participants practiced the difficult exercises more often than the easy ones during the learning period, and this difference was found to be statistically significant ($F(1,27) = 9.00$, $p < .01$). Also information presentation had an effect on the amount of practising during the learning period ($F(2,27) = 6.76$, $p < .005$). In the film clip condition, participants almost never practiced, in the picture condition they practiced for about a fifth of the exercises, while in the text condition participants practiced about half of the exercises. Post-hoc tests showed that the difference between the instruction in text and the instruction in a picture approached significance ($p = .06$). Text differed significantly from the instruction in a film clip ($p < .005$). There was no significant difference between the instruction in a picture and a film clip ($p = .43$), presumably because of the relatively high standard deviation. No significant interaction effects were found.

Execution times

The difficulty degree had a main effect on the amount of time needed to perform the exercises ($F(1,27) = 20.84$, $p < .001$). The participants required more time to execute the difficult exercises than the easy ones. There was also a main effect of information modality on the execution times ($F(2,27) = 26.78$, $p < .001$). Participants in the text condition had much longer execution times than those in the picture and film clip conditions. Participants in the picture condition needed somewhat less time for execution than the participants in the film clip condition, but the differences are relatively small and the standard deviations are relatively high. Post hoc tests indicated that there was a significant difference between the instruction in text and the instruction in a picture ($p < .001$). Also, the instruction in text significantly differed from the instruction in a film ($p < .001$). There was no significant difference between the instruction in a film clip and in a picture ($p = .35$). There was no significant interaction between difficulty degree and information modality.

Amount of correctly executed exercises

A main effect of difficulty degree on the performance was found ($F(1,27) = 11.76$, $p < .005$): as can be seen in Table 1, the participants executed on average 9.1 easy exercises correctly, as opposed to 8.3 difficult ones. There was also a main effect of information modality: the participants who watched the film clip executed the most movements correctly ($F(2,27) = 11.68$, $p < .001$). A two-way interaction between difficulty degree and information modality was found ($F(2,27) = 3.62$, $p < .05$). Post hoc tests indicated that in the film clip condition, both easy and difficult exercises are almost always executed correctly. In the text and picture condition however, easy exercises were performed correctly more often than difficult exercises. Post hoc analyses revealed that there was a significant difference between instruction via text and via film clips ($p < .001$) as well as between pictures and film clips ($p < .005$), but not between text and picture ($p = .27$).

Subjective satisfaction

[INSERT TABLE 2]

The internal consistency on the four items sets in the questionnaire was measured using Cronbach's α. Following standard practice, we qualify an α as adequate if its value was higher than .70. For the structure of the web site the α was 0.78, and for the content of the web site the α was 0.56[3]. The α for the comprehensibility of the exercises was 0.82, and for the attractiveness of the exercises 0.83.

Table 2 gives an overview of the results of the subjective satisfaction questionnaire. Information modality had no effect on the subjective satisfaction regarding the web site and the exercises. No effects were found between the three conditions for the web site ($F<1$) and for the exercises ($F<1$).

Summary

The results showed that there were significant main effects of difficulty degree and information modality on learning times, amount of practising, execution times, and learning results (in terms of correctly executed exercises). However, the subjective satisfaction of the participants regarding the web site and the exercises revealed no differences between the three information modalities. An explanation for this result could be the between-subjects design: participants only saw one realization of each exercise, and arguably could not form an informed preference for one of the three information modalities. Therefore, a second experiment was conducted to find out whether participants preferred one of these three information modalities.

[3] We report on these items separately.

Experiment II: subjective preference (forced choice) for information modalities

Method

Participants

Participants were 26 young adults, between 18 and 25 years old. Of the participants 13 were male and 13 were female. None participated in the first study.

Design

The experiment had a 3 (information modality) × 2 (difficulty degree) factorial design, with information modality and difficulty degree as within participants variable and preference as the dependent variable. The participants were randomly assigned to an experimental condition.

Stimuli

In the second experiment, participants did not have to execute the RSI exercises, but were asked which instructional format (text, picture, or film clip) they preferred for a given exercise. Eight exercises (four easy and four difficult ones) were randomly selected from the 20 exercises from experiment I. Two of these exercises (one easy and one difficult) were used to instruct the participant during the practice period, the other six were used in the actual experiment. To control for possible learning effects, the exercises were presented in four random orders, i.e., two random orders for the presentation of the information modality and two random orders for the exercises.

Procedure

Participants took part one at a time. They were invited to an experimental laboratory, and were seated behind the computer. They were told that they would receive six RSI exercises to learn in three versions (i.e., text, picture, and film clip). After learning the exercise, their task was to indicate (by forced choice) which of the three information modalities they preferred for that exercise. Following the instructions, participants could proceed with two trial exercises to make them acquainted with the stimuli and the task. For each exercise, the instructional formats were presented beneath each other (see figure 4). When the participants had observed the three instructional formats, they were instructed to fill in their preferred realization for that exercise on an answer sheet. The sequence in which the information modalities were presented on the answer sheet corresponded to the sequence in which they were presented at the web site. After the trial exercises, the experiment leader asked if the participants had any questions regarding the procedure of the experiment. If the procedure was clear participants could start the actual experiment and select their preferred mode of presentation in the same manner as during the practice session. There was no further interaction between participants and experiment leader during the experiment.

[INSERT FIGURE 4]

Results

The data were analysed with χ^2-tests to check for significant differences in participant's preferences of presentation formats for RSI exercises. Table 3 shows the result: for all exercises the majority of the participants preferred the film clip to text and picture. The overall distribution was significantly different from chance (χ^2 (2) = 81.03, p< .01). There was no effect of difficulty degree on the preference of the participants (χ^2 (5) = 1.61, n.s.). Interestingly, there were some notable differences in the distribution of preferences for the first two easy exercises. For the first easy exercise, the participants preferred the film clip and text to the instruction in a picture. For the second easy exercise, the participants preferred the film clip and picture to the instruction in text. This was a significant difference (χ^2 (5) = 12.76, p< .05).

[INSERT TABLE 3]

Summary

A second experiment was conducted to find out if participants preferred one of the three information modalities. The results of this study showed that for all exercises most participants preferred the film clips for illustrating the RSI exercises. Thus, there was a significant difference of the instructional format on the subject's preference. Difficulty degree did not influence this preference pattern.

Conclusion and discussion

In this paper, we investigated the learnability of RSI prevention exercises in different presentation formats. What effective ways of learning such exercises are, is an important research question, in view of the growing awareness of the importance of RSI prevention and the bewildering array of presentation formats for RSI prevention exercises currently being employed on the internet and in RSI prevention software.

Two experiments were performed looking at the effect of offering RSI prevention exercises in three different formats (film clips, pictures or text) on learning time, amount of practicing, execution time, learning results and subjective satisfaction. To model variation within a domain, twenty RSI prevention exercises were selected from different sources in such a way that 10 exercises were motorically easy (symmetric and simple) and 10 exercises more difficult (asymmetric or complex), and a pre-test with 9 participants revealed that there was strong connection between objective and perceived difficulty degree.[4] The results of the first experiment indeed reveal that the exercises assumed to be easy were 'easier' to perform than the assumed difficult ones, since significant main effects of difficulty degree were found on all dimensions of interest. Thus, irrespective of presentation format, the easy exercises are associated with shorter learning times, less practicing, shorter execution times and less execution errors than the difficult exercises. It is worth repeating that the summed length (in terms of frames and number of words) of the 10 easy exercises was exactly the same as the summed length of the 10 difficult ones.

Of the three presentation formats under investigation, text was expected to

[4] We redid all statistical analyses omitting the few exercises for which the subjective assessments in the pre-test did not coincide with the objective classification. This led to essentially the same results as those reported above.

impose the highest load. Overall, text indeed led to the longest learning times, which can in part be ascribed to the fact that it takes more mental effort to read a text than to watch a picture or film clip. But during the learning phase, people not only read the text, but also engage in a substantial amount of practicing which takes time as well. People in the text condition did by far the most practicing, which is consistent with the Indexical Hypothesis (e.g., Glenberg 1997): to foster understanding, participants 'translated' the textual instructions into actual movements during learning for many exercises. Participants must thus engage in fairly deep processing of the textual instructions, but arguably this is to some extent a form of germane load: the deep processing and practicing appear to be beneficial for learning. Contrary to what one might expect, the relatively long learning phase, does not lead to shorter execution times. Still it does lead to a surprisingly good learning performance. For the easy exercises, participants in the text condition make a few more errors than participants in the other two conditions, but for the difficult exercises, performance is even slightly better than for pictures, which suggests that the germane activities pay off.

Pictures were expected to impose the lowest load. The average learning times were indeed lowest in the picture condition, as were the average execution times. The pictures do not offer a complete model of the exercise, but only the stroke of the movement is depicted, with arrows indicating motion where this is applicable. The expectation was that people generally would be able to derive the complete exercise on the basis of this information. This might explain why a moderate amount of practicing took place in this condition (more than for film clips). For easy exercises, learning from pictures led to a good performance. In fact, participants made as few errors for these exercises as participants in the film clip condition. But the performance dropped for the difficult exercises, where as much errors were made as in the text condition.

Concerning the load of film clips, two contrasting hypotheses were mentioned: they might induce load because a learner has to process the continuously changing images, but they might also reduce load, freeing the learner by simply presenting a complete, physical model of the task to be carried out. It was found that film clips led to medium length learning times (between picture and text). In part this can be attributed to the fact that watching a clip takes a fixed length. But it is interesting to see that difficult exercises require longer learning times than easy ones, even though they are of the same average duration (and it is not the case that learners played difficult exercises more often), which is probably due to the higher average intrinsic load of the difficult tasks. There was virtually no practicing in the film clip condition, as expected, since the clips offer an informationally complete model of the task. Contrary to expectation, the execution times were not the shortest, which suggests that participants still had to engage in cognitive activity during the execution phase. Film clips did lead to a consistently high learning performance, both for easy and for difficult exercises. Hence, despite the apparent lack of cognitive effort during learning (no practicing), learners do construct the necessary scheme based on germane cognitive processes.

It is interesting to observe that none of the presentation formats appears to be superior on all dimensions of interest, each has some disadvantage (less efficient for learning, relatively many errors, etc.). In view of this, it is perhaps not surprising that the first study did not reveal any significant differences on the subjective preference dimension. Hence a second experiment was performed, in which participants had to select their learning preference via forced choice. They had to do so for 6 randomly selected exercises (3 easy and 3 difficult ones). No significant differences were found for the difficulty degree of the exercises, which might be explained from the fact that the participants of the second study only *observed* the exercises: they did not have to learn and execute them, and therefore might have processed these exercises on a more superficial level. Interestingly, a general and consistent preference for film clips was found (contrary the results of the first experiment). What causes this apparent discrepancy between the effectiveness of information modalities and the subjective learning preference is not entirely clear. Part of the explanation may be that the film clips are the most 'visually appealing'. In addition, it may be that participants of the second experiment recognize that film clips offer a complete action representation, but do not realize that learning from text or a picture may lead to good results as well (and perhaps even quicker than for film clips, see above). The relation between perceived and actual usefulness of different information formats is an interesting one, which we would like to pursue in more detail in future research.

RSI prevention exercises offer a new and ecologically valid learning situation with a number of interesting properties. These exercises are quite brief, and arguably relatively easy to learn. A downside of this is that with respect to learning performance (number of errors) there may be a ceiling effect, in that easy exercises are learned very well for all three modalities. It would be interesting to redo the experiment with more complex RSI related tasks, and see whether this would lead to more differentiation between the different modalities where errors are concerned.

While we did our best to make sure that the exercises in the three conditions offer comparable information, following the recommendations of Tversky et al. (2002), it turned out that this was not always straightforward. While a static picture combined with an arrow indicating direction or motion can be very informative, it does not make the entire intended movement explicit as a dynamic picture does. In the former, but not in the latter case, the learner has to infer the full movement, which may lead to errors, especially for the difficult exercises as we have seen. Still, it is interesting to see that the efficiency of pictures (learning and execution times) is higher than that of the other two modalities, and leads to nearly optimal results for easy exercises. This indicates that a particularly efficient method for illustrating more complex exercises might be via a series of pictures, depicting key stages of the procedure. One would expect that this could lead to both a high efficiency and good learning results, for easy as well as for difficult exercises. We intend to explore this possibility in future research.

In a somewhat similar vein, we found that certain RSI exercises seem to be inherently easier to represent than others, and especially that this ease-of-representation may vary across different presentation formats. Some movements can be very concisely described in language, because the entire movement has been "coded" in a fixed expression (e.g., “make fists”), whereas other movements can be rather cumbersome to describe. Also expressing how a particular movement ‘feels’ (“spread your fingers until a mild stretch between the fingers is felt”) is obviously easier in language than in static or dynamic visuals. For such exercises, a textual presentation might have an added value over other presentation formats. It would be interesting to systematically vary the presence or absence of linguistic short cuts (describing complex movements in a few words) and investigate how this influences the effectiveness of textual instructions.

Acknowledgements

The current research is part of the Interactive Multimodal Information eXtraction (IMIX) project, which is funded by the Netherlands Organisation for Scientific Research (NWO). Thanks are due to Pashiera Barkhuysen and Lennard van der Laar for their help in setting up the experiment, and to Fons Maes and Nicole Ummelen for useful comments on an earlier version of this paper. We greatly benefited from the constructive comments of three anonymous reviewers.

Appendix A Questionnaire experiment 1.

English translation of Dutch originals:

The structure of the web site is:

- Orderly / Disorderly
- Clear / Unclear
- Easy to access / Uneasy to access
- Easy to see through / Uneasy to see through

The content of the website is:

- Informative / Uninformative
- Clear / Unclear
- Comprehensible / Incomprehensible
- Understandable / Unintelligible

The comprehensibility of the exercises is:

- Easy / Difficult
- Simple / Hard
- Clear / Unclear
- Unambiguous / Ambiguous

The attractiveness of the exercises is:

- Varied / Unvaried
- Interesting / Uninteresting
- Appealing/ Unappealing
- Fascinating / Tiresome

References

Baddely, A. (1992). Working memory. *Science, 255,* 556-559.

Balci, R., Aghazadeh, F. (2003). The effect of work-rest schedules and type of task on the discomfort and performance of VDT users. *Ergonomics, 46*, 455-465.

Betrancourt, M., & Tversky, B. (2000). Effects of computer animation on user's performance: a review. *Le travail humain*, *63*, 311-329.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisation. *Learning and Instruction*, *14*, 325-341.

Brünken, R., Plass, L, and Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning, *Educational Psychologist, 38,* 53-61.

Carroll, L. & Wiebe, E. N. (2004). Static versus dynamic presentation of procedural instruction: Investigating the efficacy of video-based delivery. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Santa Monica, CA: HFES.

Chi, M., Glaser, R., Rees, E. (1982). Expertise in problem solving. In Sternberg, R. (ed.), *Advances in the Psychology of Human Intelligence*, Erlbaum, Hillsdale NJ, pp. 7-75.

Cheng, P. (2002). Electrifying diagrams for learning: principles for complex representational systems. *Cognitive Science, 26*(6), 658-736

Glenberg, A. (1997). What memory is for. *Behavioral and Brain Sciences, 20*, 1-19.

Glenberg, A., & Robertson, D. (1999). Indexical Understanding of Instructions. *Discourse Processes, 28*, 1-26.

Hegarty, M. (2004). Dynamic visualizations and learning: getting to the difficult questions. *Learning and Instruction, 14*, 343-351.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M.R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207-227). The Hague: Mouton.

Kirby, J., Moore, P. and Schofield, N. (1988). Verbal and visual learning styles, *Contemporary Educational Psychology, 13,* 169-184.

Larkin, J. & Simon H. (1987), Why a diagram is (sometimes) worth a thousand worths, *Cognitive Science*, *11,* 65-99.

Leutner, D. and Plass, J. (1998). Measuring learning styles with questionnaires versus direct observation of preferential choice behavior in authentic learning situations: The Visualizer/Verbalizer Behavior Observation Scale (VV-BOS), *Computers in Human Behavior, 14,* 543-557.

Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction*, *13*, 177-189.

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14*, 257-274.

Marcus, N., Cooper, M. & Sweller, J. (1996). Understanding Instructions, *Journal of Educational Psychology, 88*, 49-63.

Mayer, R. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81*, 240-246.

Mayer, R. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mayer, R., & Gallini,. J. (1990). When is an illustration worth a thousand words? *Journal of Educational Psychology, 82*, 715-726,

Mayer, R., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*, 107-119.

Mayer, R. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13*, 125-139.

Mclean, L., Tingley, M., Scott, R., & Rickards, J. (2001). Computer terminal work and the benefit of micro-breaks. *Applied Ergonomics, 32*, 225-237.

van Merriënboer, J., Schuurman, J., de Croock, M., Paas, F. (2002). Redirecting leaners' attention during training: effects on cognitive load, transfer test, performance and training efficiency. *Learning and Instruction, 12,* 11-37.

Michas, I., & Berry, D. (2000). Learning a procedural task: effectiveness of multimedia presentations. *Applied Cognitive Psychology, 14,* 555-575.

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Nielsen, J. (1993). *Usability engineering*. New York: Academic press.

Park, O., & Hopkins, R. (1993). Instructional conditions for using dynamic visual displays: a review. *Journal of Instructional Science, 21*, 427-449.

Park, O. & Gittelman, S. (1995). Dynamic Characteristics of mental models and dynamic visual displays. *Journal of Instructional Science*, *23*, 303-320.

Pane, J., Corbett, A., & John, B. (1996). Assessing dynamics in computer-based instruction. In: *Proceedings of the ACM Conference on Human factors in Computing Systems*, Vancouver, pp. 797-804.

Perez, E., & White, M. (1985). Student evaluation of motivational and learning attributes of microcomputer software. *Journal of Computer-Based Instruction, 12*, 39-43.

Ploetzner, R., & Lowe, R. (2004). Dynamic visualisations and learning. *Learning and Instruction, 14*, 235-240.

Reed, S. (1985). Effects of computer graphics on improving estimates to algebra word problems. *Journal of Educational Psychology, 77*, 285-298.

Rieber, L. (1991). Animation, incidental learning, and continuing motivation. *Journal of Educational Psychology, 83,* 318-328.

Schnotz, W., Böckheler, J., Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures, *European Journal of Psychology of Eduction*, 14, 245-265.

Schnotz, W., Rasch, T. (2005), Enabling, facilitating. and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning, *Educational Technology Research & Development, 53,* 47-58.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction, 14*, 293-305.

Stone, W. (1983). Repetitive strain injuries. *Medical Journal of Australia 10*, 616-618.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition & Instruction, 8,* 351-362.

Sweller, J., van Merriënboer, J., Paas, F. (1998), Cognitive Architecture and Instructional Design, *Educational Psychology Review*, 10, 251-296.

Tversky, B., Zacks, J., Lee, P., & Heiser, J. (2000). Lines, blobs, crosses, and arrows: diagrammatic communication with schematic figures. In: *Theory and application of diagrams*, M. Anderson, P. Cheng, V. Haarslev (eds.), Berlin: Springer, pp. 221-230.

Tversky, B., Morrison, J., & Betrancourt, M. (2002). Animation; can it facilitate? *Int J. Human-Computer Studies, 57*, 247-262.

Weidenmann, B. (1988). When good pictures fail. An information processing approach to the effect of illustrations. In: *Knowledge acquisition from text and pictures*, H. Mandl, J. Levin (eds.), Amsterdam, Elsevier, pp. 157-171.

Weiss, R., Knolton, D., & Morrison, G. (2002). Principles for using animation in computer-based instruction: theoretical heuristics for effective design. *Computers in Human Behavior*, *18,* 465-477.

Williams, T., Smith, L, & Herrick, R. (1989). Exercise as a prophylactic device against carpal tunnel syndrome. *Proceedings of the 33rd Annual Meeting of the Human Factors Society*, Santa Monica, CA, pp. 723-727.

Fig. 1
Three different web sites, which use different formats (text, picture and text, and animation) to illustrate RSI prevention exercises

(http://naturalhealthcare.ca/rsi_saver_exercises.phtml;
http://web.mit.edu/atic/www/disabilities/rsi/exercises.html;
http://busy-bee.net/rsi/)

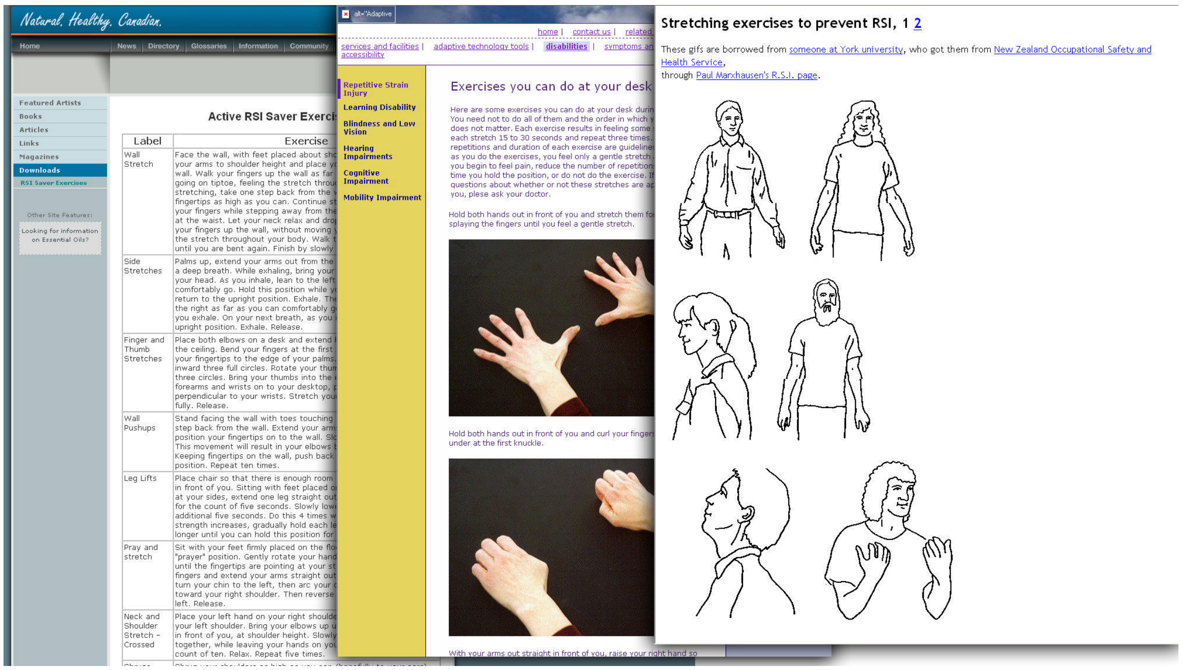
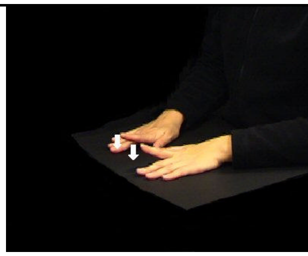


Fig. 2

Two typical exercises: one easy exercise and one difficult exercise

Easy exercise	
	Hold your hands in front of you with your palms facing downwards. Lift both your index fingers from the knuckles. Next, gently drop your index fingers. (*English translation of Dutch original, in Dutch this exercise contains 29 words*)

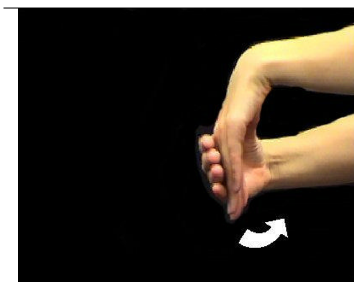
Difficult exercise	
	Hold your left arm in front of you and drop the left hand down bending at the wrist. Place your right hand on the knuckles of the left hand. Press your right hand towards. (*English translation of Dutch original, in Dutch this exercise contains 30 words*)

Fig. 3

Procedure of experiment I

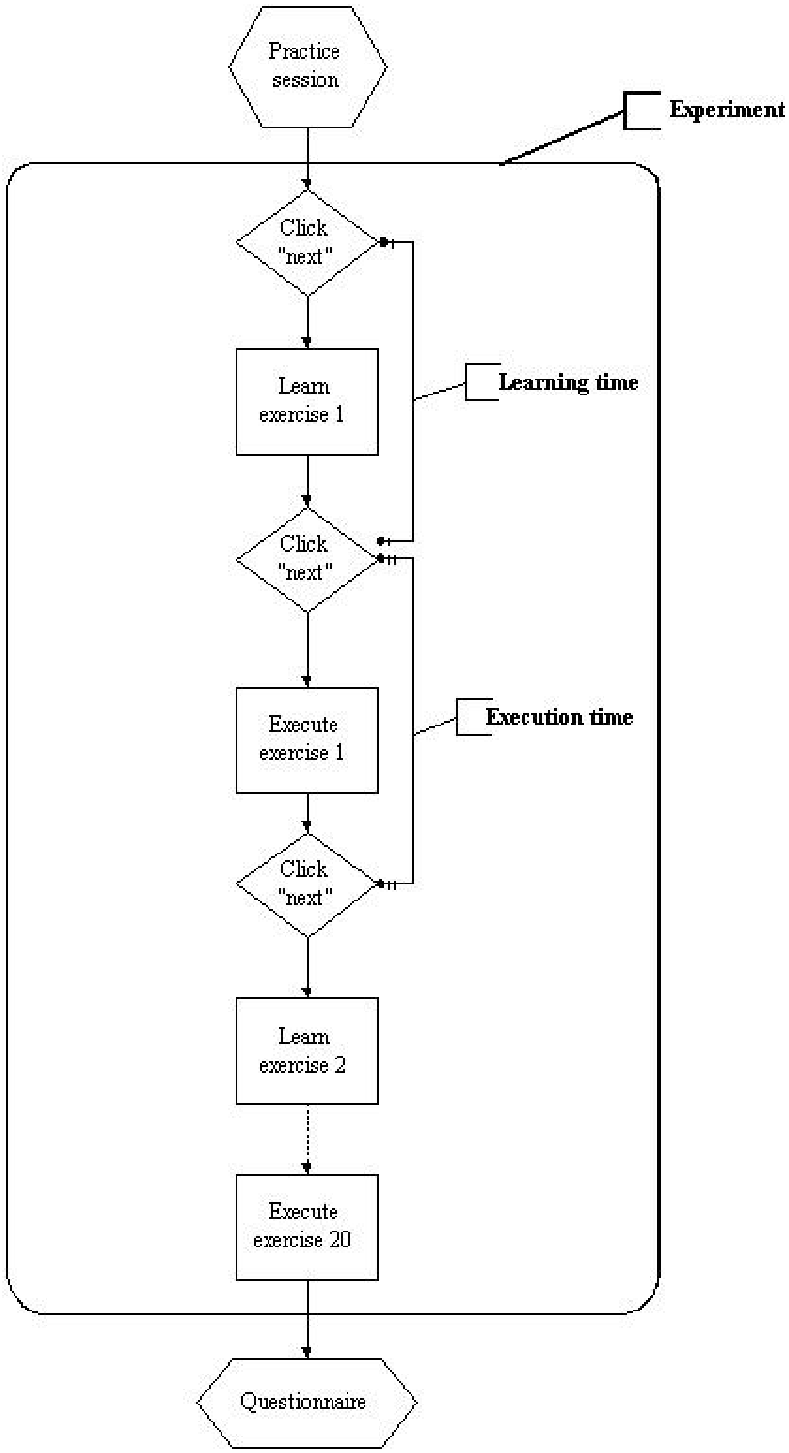


Table 1

Mean total time taken to learn and execute the exercises, average number of exercises for which participants practised during the learning period and the number of correctly executed exercises as a function of difficulty degree for the three information modalities of interest (standard deviations in parenthesis)

Factor	Difficulty degree	Text		Picture		Film clip		Average	
Learning time	Easy	96.30	(45.04)	55.10	(17.51)	76.10	(17.07)	75.83	(33.28)
	Difficult	108.10	(40.64)	69.80	(29.04)	91.20	(15.42)	89.70	(33.20)
Practicing	Easy	4.70	(4.79)	1.50	(1.84)	0.00	(0.00)	2.07	(3.48)
	Difficult	5.30	(4.62)	2.20	(2.30)	0.40	(0.70)	2.63	(3.56)
Execution time	Easy	164.20	(48.86)	53.80	(15.48)	77.90	(13.42)	98.63	(56.52)
	Difficult	196.10	(71.28)	77.00	(27.95)	99.60	(31.41)	124.23	(69.89)
Correctly executed exercises	Easy	8.40	(0.70)	9.20	(0.63)	9.60	(0.70)	9.10	(0.84)
	Difficult	7.90	(1.10)	7.60	(1.35)	9.40	(0.70)	8.27	(1.44)

Table 2

Mean results of the subjective satisfaction questionnaire regarding the structure of the web site, the comprehensibility and attractiveness of the exercises, and the content of the web site's exercises in relation to the 3 experimental conditions (scores range from 1 = "very positive" to 7 = "very negative"; standard deviations in parenthesis). Since the α for Content was below the threshold, we report on the four components separately.

Factor	**Subjective satisfaction regarding**	**Text**	**Picture**	**Film clip**
Web site	Structure	2.25 (0.99)	1.85 (0.58)	2.23 (1.11)
	Content			
	Informative	4.30 (1.95)	3.40 (1.51)	4.20 (1.40)
	Clear	2.50 (1.18)	2.70 (1.49)	2.70 (1.06)
	Comprehensible	2.90 (1.29)	3.30 (1.77)	2.30 (1.43)
	Understandable	3.70 (1.70)	3.00 (2.00)	2.40 (1.42)
Exercises	Comprehensibility	3.80 (1.05)	3.93 (1.32)	2.80 (1.44)
	Attractiveness	3.57 (1.08)	4.07 (1.19)	3.50 (.93)

Fig. 4

Example of how the exercises were presented in the web sites during the second experiment.

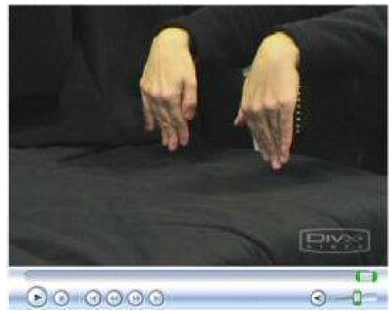
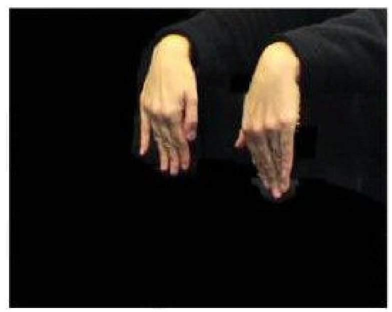
A. Film clip	
B. Picture	
C. Text	Put your hands in front of you, with your palms facing the table. Now gently stretch your wrists down until your fingertips point downward.

Table 3

The distribution of the answers on the questions regarding the preference in relation to the easy exercises and the information modalities

	Text	Picture	Film clip	Total
Easy 1	10	3	13	26
Easy 2	1	12	13	26
Easy 3	3	0	23	26
Difficult 1	6	1	19	26
Difficult 2	3	4	19	26
Difficult 3	3	5	18	26
Total	26	25	105	156