British Journal of Educational Psychology (2003), 73, 489–505

www.bps.org.uk

The efficiency of multimedia learning into old age

Pascal W.M. Van Gerven[1]*, Fred Paas[2], Jeroen J.G. Van Merriënboer[2], Maaike Hendriks[2] and Henk G. Schmidt[1]

[1]Institute of Psychology, Erasmus University Rotterdam, The Netherlands
[2]Educational Technology Expertise Centre (OTEC), Open University of The Netherlands

Background. On the basis of a multimodal model of working memory, cognitive load theory predicts that a multimedia-based instructional format leads to a better acquisition of complex subject matter than a purely visual instructional format.

Aims. This study investigated the extent to which age and instructional format had an impact on training efficiency among both young and old adults. It was hypothesised that studying worked examples that are presented as a narrated animation (multimedia condition) is a more efficient means of complex skill training than studying visually presented worked examples (unimodal condition) and solving conventional problems. Furthermore, it was hypothesised that multimedia-based worked examples are especially helpful for elderly learners, who have to deal with a general decline of working-memory resources, because they address both mode-specific working-memory stores.

Sample. The sample consisted of 60 young (mean age = 15.98 years) and 60 old adults (mean age = 64.48 years).

Methods. Participants of both age groups were trained in either a conventional, a unimodal, or a multimedia condition. Subsequently, they had to solve a series of test problems. Dependent variables were perceived cognitive load during the training, performance on the test, and efficiency in terms of the ratio between these two variables.

Results. Results showed that for both age groups multimedia-based worked examples were more efficient than the other training formats in that less cognitive load led to at least an equal performance level.

Conclusion. Although no difference in the beneficial effect of multimedia learning was found between the age groups, multimedia-based instructions seem promising for the elderly.

* *Requests for reprints should be addressed to Pascal W.M. Van Gerven, Faculty of Psychology, Department of Neurocognition, Maastricht University, PO Box 616, 6220 MD Maastricht, The Netherlands (e-mail: P.vanGerven@psychology. unimaas.nl).*

The strong increase of the ageing population in Western societies has provoked a true avalanche of research aimed at numerous aspects of both physical and psychological ageing. For that matter, educational research has surely not remained an unexplored realm. Although ageing is often associated with infirmity, it can be expected that the proportion of independent, active, and eager-to-learn elderly people is growing just as well. It is therefore not surprising that educational research has put forward *lifelong learning* as one of its central objectives. Lifelong learning can be achieved by an instructional intervention that efficiently addresses the available cognitive capacity. For that purpose, principles of instructional design have to be combined with outcomes of cognitive ageing research. It is peculiar, however, that little of the work done in this field is actually based on such an integrative theory (exceptions are Czaja, 1996; Morrell & Echt, 1996, 1997; Morrell & Park 1993). Van Gerven, Paas, Van Merriënboer, and Schmidt (2000) propose a framework in which cognitive load theory (Sweller, 1988; Sweller, Van Merriënboer, & Paas, 1998) is merged with general views on cognitive ageing (e.g., Hasher & Zacks, 1988; Salthouse 1996; Salthouse & Babcock, 1991).

Cognitive load theory (CLT) is an instructional theory that starts from the idea that human working memory is limited with respect to the amount of information it can hold and the number of operations it can perform on that information. CLT furthermore assumes that working memory plays an essential role in skill learning, because it serves as a device for transforming instructional information into useful *cognitive schemata*. Cognitive schemata are information structures in long-term memory that enable someone to solve a certain category of problems and at the same time save working memory capacity by chunking information elements and production rules into a whole. Moreover, cognitive schemata facilitate transfer performance (e.g., Paas & Van Merriënboer, 1994a), which is the use of acquired knowledge in problem situations that more or less deviate from problem situations that were encountered during training. The central idea of CLT is that a learner should be stimulated to use his or her limited working memory efficiently, especially when acquiring a complex skill. That is, in order to achieve optimal learning, working memory should be occupied by a minimum number of irrelevant operations and a maximum number of relevant operations. This can be achieved by a proper instructional design for which CLT promotes a series of guidelines (see Sweller *et al.*, 1998, for an overview).

A theory concerned with cognitive limitations may especially be applicable to the acquisition of complex cognitive skills on behalf of learners whose cognitive capacity is particularly limited, such as the elderly. Healthy elderly people have to cope with at least three cognitive declines. First, there is a decrease of working-memory capacity (e.g., Salthouse & Babcock, 1991; Wingfield, Stine, Lahar, & Aberdeen, 1988), which is the ability to store and manipulate a relatively small amount of information for a relatively short period. Second, there is a decline of the rate at which information is activated and processed (Myerson, Hale, Wagstaff, Poon, & Smith, 1990; Salthouse, 1996). This reduced cognitive speed can lead to poor comprehension and learning as a result of what Salthouse (1996) calls the 'simultaneity mechanism'. If mutually dependent information elements are not activated quickly enough, activation of one element can have dissipated before activation of the other has reached a certain threshold. A third decline concerns the failure to suppress information that is not relevant for a task at the cost of capacity available for relevant processing (e.g., Allen, Madden, Groth, & Crozier, 1992; Hasher & Zacks, 1988; McDowd, 1997).

According to Van Gerven *et al.* (2000), CLT offers the principles to deal optimally with these three declines. The decreased-capacity problem can be largely compensated by preventing the learner from getting engaged in task-irrelevant processes. The same holds for the decline of cognitive speed: the probability of attaining simultaneous activation of interdependent information elements can be increased by merely allowing task-relevant information to become active in working memory. Problems with suppressing irrelevant information, finally, can be avoided by presenting a minimal number of redundant stimuli. These principles may seem rather obvious. In conventional instructional settings, however, they are hardly taken into account.

This is especially true for the practice phase of an instruction. Complex skills are usually trained by solving a relatively large number of practice problems. Problem solving as a means of training has a significant drawback, however. It imposes a high cognitive load on the learner, which is, as we will see, especially problematic in an early stage of learning. This high level of cognitive load is caused by a commonly used strategy for solving relatively unfamiliar problems, namely *means-ends analysis*. Means-ends analyses take place in two phases. First, a range of goals and subgoals is set, working backward from the main goal. When a level of attainable subgoals is reached, the problem solver works forward from these subgoals toward the main goal. It may be clear that keeping in mind the whole hierarchy of goals and subgoals as well as finding the right operators to attain them puts a considerable burden on working memory. When cognitive resources are sufficient, means-ends analysis may be regarded as an effective problem-solving strategy. When resources are limited, however, problem solving through means-ends analysis is hardly effective.

In both cases, the learning result is far from optimal, because too many resources are used for irrelevant mental operations (Sweller, 1988). The irrelevant or, in terms of CLT, extraneous part of means-ends analysis concerns the backward-processing or goal-setting phase, because it does not contribute to the acquisition of cognitive schemata. For the construction of cognitive schemata, problem states have to be associated with useful operators. According to CLT, this process should not be obstructed by any irrelevant activity in working memory.

Instead, cognitive resources should be saved by preventing the learner from using 'weak' solving strategies, such as means-ends analyses, which require backward processing. A CLT-based alternative in this respect is training with *goal-free problems*. In goal-free problems no clear-cut goal state is requested. In the domain of mathematics, for instance, the learner could be asked to solve an equation for as many variables as possible instead of one particular variable. In this way, forward rather than backward processing is encouraged, which is beneficial for the learning process. Training with goal-free problems was found to result in better performance than conventional problem solving in young adults (Bobis, Sweller, & Cooper, 1994; Owen & Sweller, 1985; Sweller & Levine, 1982) and recently also in older adults (Paas, Camp, & Rikers, 2001).

Another important CLT-based alternative for conventional problem solving is the use of *worked examples*. Worked examples are problems accompanied by the subsequent solution steps that lead to a particular goal state. According to CLT, worked examples impose relatively little cognitive load on the learner by directly focusing attention on problem states and their corresponding solution operators. This is especially beneficial in an early stage of learning, in which specific cognitive schemata do not yet exist. Without cognitive schemata, the learner will rely on conscious processing. Conscious processing of complex material is rather demanding in terms of cognitive resources, so that overload is likely to occur. Provided that the learner is optimally motivated, studying worked examples quickly leads to the construction of rudimentary, but indispensable schemata. This saves processing capacity, which can be used for transforming these undeveloped schemata into more elaborate ones. Superior performance using worked examples as opposed to conventional problems was repeatedly demonstrated (Paas, 1992; Paas & Van Merriënboer, 1994b; Sweller & Cooper, 1985; Ward & Sweller, 1990). Van Gerven, Paas, Van Merriënboer, and Schmidt (2002) found worked examples to be a more efficient training method for elderly learners in terms of invested mental effort.

There are numerous ways to design worked examples. The most straightforward one is a combination of images and explanatory text. These two sources of information are often physically separated (like the figures and text in this article). Since images and text are often interdependent in that they cannot be understood on their own, they have to be mentally integrated, which not only requires visual search but also extraneous activity in working memory. This *split attention* phenomenon can largely be prevented by physically integrating the text with the visual material, for instance by imposing text fragments on those parts of the image that they refer to. Research has shown that this integrated format leads to better performance than conventional non-integrated formats (Bobis, Sweller, & Cooper, 1993; Chandler & Sweller, 1992; Sweller & Chandler, 1994; Sweller, Chandler, Tierney, & Cooper, 1990).

The modality effect

An alternative approach to deal with split attention is to address different sensory modalities during training. More specifically, visual material can be presented through the visual channel and explanatory text via the auditory channel, which can easily be achieved with modern multimedia techniques. There is a long tradition of research exploring the effects of information presented via different modal channels (see Penney, 1989, for a review). Much of the theory in this area is based on the idea that the visual and auditory channels correspond to partly independent working-memory stores (e.g., Baddeley, 1992). An additional idea is the concept of 'dual coding', which emphasises the benefits of encoding information in both visual and auditory form (e.g., Clark & Paivio, 1991). CLT has adopted these views and claims that addressing different sensory modalities reduces extraneous cognitive load in that it decreases visual search and the need for mental integration. Cognitive load is furthermore reduced, because incoming information is distributed over different limited-capacity stores. The overall reduction of instruction-related cognitive load increases the likelihood of grasping complex subject matter.

Although some researchers are sceptical towards the value of 'multi-representational' learning environments (e.g., Ainsworth, Wood, & O'Malley, 1998), evidence in favour of multimodal instructional environments is overwhelming. In the context of CLT, the modality effect was first demonstrated by Mousavi, Low, and Sweller (1995) with high school students in the domain of geometry. Subsequently, the effect was obtained by Mayer, Bove, Bryman, Mars, and Tapangco (1996) in the domain of meteorology, by Tindall-Ford, Chandler, and Sweller (1997) in the domain of electrical engineering, and most recently by Kalyuga, Chandler, and Sweller (1999, 2000) in other technical domains. A related set of studies by Mayer and associates (Mayer & Anderson, 1991; Mayer & Moreno, 1998; Moreno & Mayer, 1999) stresses the importance of *temporal contiguity* in the design of multimedia training. Temporal contiguity refers to the synchronisation of pictorial and auditory information. Overall, it was found that adequate timing of narrated text plays an essential role in comprehension and recall of multimedia-based training material (see Mayer & Moreno, 2002, for an overview).

The aim of the present study was to investigate the modality effect with animated worked examples in both young and older adults by combining the modality effect, which springs from cognitive load theory, and the contiguity effect, which arises from Mayer's cognitive theory of multimedia learning. This makes it a sequel to an earlier study (Van Gerven *et al.*, 2002), which investigated the relative efficiency of purely visual worked examples in the domain of Luchins's (1942) 'water-jug problem'. It was found that worked examples, which were presented as a set of pictures representing the subsequent solving steps of a problem, were more efficient relative to conventional problems in that less training time and less invested cognitive load led to a comparable performance level. For a couple of reasons these *unimodal* worked examples could still have caused a substantial extraneous load, however. First, they required the learners to combine textual and pictorial information within a single picture. Second, they required integrating the information of up to six pictures within an example. Unnecessary search and integration processes could thus have been the result. A final source of extraneous load could have been the repetition of visual information, which is intrinsic to a step-by-step representation and is likely to cause unnecessary distraction.

A *multimedia* version of these worked examples, in which pictorial information is presented in an animated and the text in an auditory form, will not have all the aforementioned disadvantages. For elderly learners multimedia-based worked examples are expected to have additional benefits. Optimal contiguity of visual and auditory information will minimise the probability of simultaneity problems (Salthouse, 1996). In addition, animations will avoid the repetition of invariable pictorial elements, which are likely to distract attention and cause search problems (McDowd, 1997). These notions bring us to formulate the following two hypotheses.

On the basis of Sweller *et al.*'s cognitive load theory and Mayer's cognitive theory of multimedia learning, Hypothesis 1 states that worked examples presented in a multimedia format are more efficient than both conventional problems and unimodally presented worked examples in that at least an equal level of performance can be obtained with less effort. Following Van Gerven *et al.*'s (2000) framework, Hypothesis 2 states that elderly learners will profit more from multimedia-based worked examples relative to unimodal worked examples and conventional problems than their young counterparts. Moreover, both hypotheses imply that unimodal worked examples are superior to conventional problems, which was already demonstrated by Van Gerven *et al.* (2002).

Method

Participants

Sixty secondary school students (24 men and 36 women; mean age = 15.98 years, *SD* = 0.77) were recruited from a high school in Heerlen. Sixty elderly participants (34 men and 26 women; mean age = 64.48 years, *SD* = 4.92) were recruited in Maastricht and the surrounding area via advertisements in local newspapers. All participants were in good health and had normal or corrected-to-normal vision. They received 40 Dutch guilders (about €18) payment for their participation.

On a shortened version of the Groningen Intelligence Test (GIT; Luteijn & Van der

Ploeg, 1983), the elderly participants ($M = 126.83$, $SD = 8.38$) scored about 11 IQ points higher than their young counterparts ($M = 115.68$, $SD = 8.73$), $t(118) = -7.14$, $p < .001$. The GIT is a commonly used Dutch estimate of formal IQ. The shortened version includes four subtasks (out of nine), which are supposed to give the best approximation of the full-scale IQ: vocabulary [woordenlijst], mental rotation [legkaart], mental arithmetic [cijferen], and analogies [woordmatrijzen]. The 95%-confidence interval of the shortened GIT is $\pm$ 10 points.

Working memory abilities were assessed with Salthouse and Babcock's (1991) computation-span test, which is a variant of the widely used reading-span test. During the computation-span test, the participant has to work out simple adding and subtraction sums, which were presented verbally (e.g., '7 plus 8', '9 minus 4', etc.). At the same time, the last digit of each sum (e.g., 8, 4, etc.) has to be memorised and reproduced after a series. A series comprises up to 7 sums and there are three trials for each. The computation-span score is determined by the maximum number of digits recalled in at least two trials. As expected, the computation span of the older participants ($M = 3.60$, $SD = 1.24$) was significantly lower than the computation span of the young participants ($M = 4.62$, $SD = 1.32$), $t(118) = 4.36$, $p < .001$.

Apparatus and materials

Apparatus

The experiment was run on an IBM-compatible computer. Interaction with the computer took place via a Philips 17-inch monitor with an integrated touch screen (Elo TouchSystems, Fremont, CA). The software controlling the experiment was programmed in Authorware 3.5 (1996).

Subjective cognitive-load scale

A 9-point symmetrical category scale was used as a subjective cognitive-load (SCL) measure (after Paas, Van Merriënboer, & Adam, 1994). The scale was presented on the touch screen as a slide control and was accompanied by the phrase: *I have experienced the foregoing as: ...* The scale had three labels, ranging from *not difficult at all* (1), via *averagely difficult* (5), to *very difficult* (9). Intermediate values had no labels. Paas (1992) found an internal consistency coefficient (Cronbach's α) of .90 using a comparable scale. Paas and Van Merriënboer (1994b) evaluated the scale as a highly reliable and sensitive instrument for the assessment of cognitive load ($\alpha = .82$). Although more objective cognitive-load measures are currently explored within CLT research (see Paas, Tuovinen, Tabbers, & Van Gerven, 2003, for an overview), the subjective scaling technique is still the most commonly used measure within the field.

Water-jug problem

A computerised version of Luchins's (1942) water-jug problem was selected as the task domain. The goal of the water-jug problem is to acquire a certain amount of water that had to end up in a target jug. For that purpose, the participants had an infinite amount of water at their disposal and a set of so-called 'working jugs' of different sizes. The working jugs could be filled and poured into each other, but they could never overflow. That is, when a bigger jug was poured into a smaller one, pouring stopped once the smaller one was full. Thus, a residual remained in the bigger jug. The goal could be reached by adequately pouring working jugs into each other and strategically making use of the residuals (for more details, see Van Gerven *et al.*, 2002).

Interface

The interface of the water-jug problem is depicted in Figure 1. The display contains several objects: a tap in the upper left corner, a button labelled *Empty* in the upper right corner, a set of working jugs in the lower left corner, and a target jug in the lower right corner. Contents are indicated inside the jugs, capacities beneath the jugs. Touching meaningful combinations of two objects resulted in three possible operations. Touching the tap in combination with one of the working jugs resulted in filling that particular working jug. Touching the *Empty* button and one of the working jugs or the target jug resulted in emptying that jug. Touching a working jug and then touching another working jug or the target jug resulted in pouring water from the former to the latter.

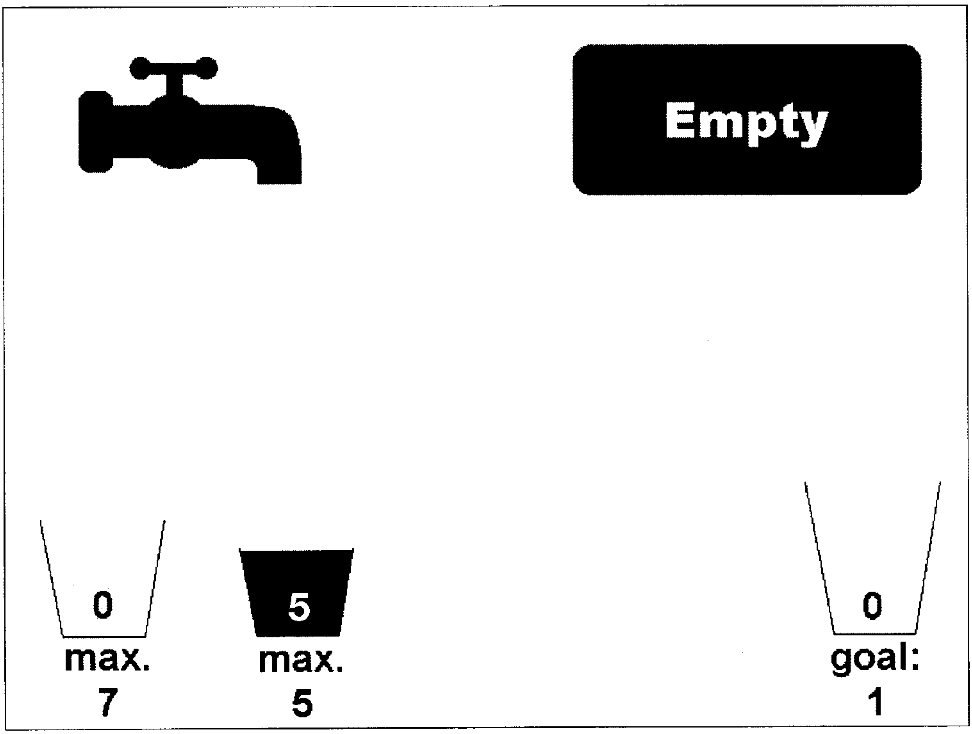

Figure 1. Interface of the water-jug task

Training problems and conditions

Four basic water-jug problems were selected as training problems. These were presented in three different forms, corresponding to the three experimental training conditions. In the conventional problems (CP) condition, the participant was presented with the initial state of a problem, after which the goal state had to be attained independently. In the unimodal worked examples (UWE) condition, the training problems were presented in a cartoon-like fashion: the initial state, the intermediate states, and the goal state were represented as a sequence of pictures. Explanatory text was imposed on the images. The CP and the UWE conditions were identical to the conventional-problems and worked-examples condition in the Van Gerven *et al.* (2002) study. In the multimedia-based worked examples condition (MWE), finally, the training problems were presented in an animated form, which was accompanied by a narrated explanation. The contents of this verbal explanation resembled the text in the UWE condition as much as possible. The animations and corresponding narration were recorded with Lotus ScreenCam (1997). In accordance with Moreno and Mayer's (1999) contiguity principle the events in the animations were temporally synchronised with the narrative's content.

Test problems

The test consisted of 12 water-jug problems. The problems were sequenced such that the test gradually increased in difficulty. This was done to prevent discouragement in the elderly participants.

Design and data collection

Participants of both age groups were randomly assigned to one of the three experimental training conditions ($n = 20$). This yielded a basic 2 (age group) × 3 (training format) between-groups design, with age group comprising the levels young and old, and training format comprising the levels conventional problems (CP), unimodal worked examples (UWE), and multimedia-based worked examples (MWE). IQ scores were considered as a covariate.

The core dependent variable was training efficiency, which reflects the ratio between subjective cognitive load (SCL) and performance. A training format can be considered inefficient if SCL during training is relatively high and performance during the test is relatively low. In contrast, a training format can be considered efficient if SCL is relatively low and performance is relatively high. Paas and Van Merriënboer's (1993) procedure was followed to convert SCL and performance scores into efficiency scores. In this procedure, a training format is represented as a dot in a co-ordinate system with the mean z score of SCL on the x-axis and the mean z score of performance on the y-axis (see Figure 3a). The z scores are calculated with the grand means and standard deviations of an age group. The training efficiency (E) score is determined by the perpendicular distance between a dot and the diagonal $E = 0$, where SCL and performance are in balance. Thus, E is calculated as follows:

$$E = \frac{^{Z}\text{Performance}^{-Z}\text{SCL}}{\sqrt{2}}$$

Other dependent variables were subjective cognitive load (SCL) during training, training time (in seconds), and performance (proportion of solved problems).

Procedure

The experiment was administered in individual sessions. A session consisted of a test block, comprising the intelligence and computation-span test, and the experiment itself. The order of these two parts was counterbalanced to control for fatigue: one half of the participants first went through the test block and then through the experiment, the other half first went through the experiment and then through the test block. The experiment consisted of a general introduction, presented as text and illustrations on the computer screen, a practice session to get acquainted with the interface, a training session (depending on the training format condition), and a final test. During the experiment, the dependent variables training time, SCL, and performance were recorded.

Results

The data were analysed with a 2 × 3 (age group × training format) between-groups overall ANCOVA, entering IQ as a covariate to control for the influence of intellectual ability on the effectiveness of the training formats. In addition, all pairwise comparisons (i.e., the contrasts CP-MWE, UWE-MWE, and CP-UWE) were analysed with separate 2 × 2 between-groups ANCOVAs, applying a Bonferroni correction for alpha: $\alpha = .05/3 = .017$. The means and standard deviations of the dependent variables are given in Table 1. As described in the Method section, training efficiency is a composite of SCL and performance. Before reporting the ANCOVA results of training efficiency, an account is given of these constituent variables. Finally, the results of training time are presented.

Table 1. Means and standard deviations per age group and training format

Age group		Training format					
		CP		UWE		MWE	
		M	*SD*	*M*	*SD*	*M*	*SD*
Young	SCL	2.54	1.16	1.98	1.00	1.94	1.46
(n = 60)	Performance	.90	.10	.84	.18	.91	.14
	Training efficiency	−0.10	0.82	−0.11	1.36	0.29	1.19
	Training time	49.70	25.01	30.33	10.63	70.88	3.91
Old	SCL	4.04	1.81	2.58	1.51	2.01	1.26
(n = 60)	Performance	.46	.26	.53	.20	.54	.22
	Training efficiency	−0.62	1.50	0.17	0.92	0.45	0.98
	Training time	123.06	68.48	110.60	74.23	77.88	10.58

Note. Performance is indicated in the proportion of solved problems. Training time is indicated in seconds.

Subjective cognitive load

With respect to SCL, no main effect of IQ was found, $F(1, 111) = 1.89$, $MSE = 1.92$, $p > .1$. Furthermore, no IQ × training format interaction could be detected, $F(2, 111) = 0.94$, $p > .3$. This suggests that IQ has no effect on perceived training load and on the impact of training format. It was therefore excluded from the rest of the overall analysis. Subsequently, a main effect of age group was found, $F(1, 114) = 8.15$, $MSE = 1.94$, $p < .01$, indicating that the elderly participants experienced higher levels of cognitive load during training than the young (see Figure 2). A main effect of training format, $F(2,$

114) = 9.78, $p < .001$, signifies that the level of SCL during training with worked examples is lower than the level of SCL in the conventional-problems condition. There is no significant age group × training format interaction, $F(2, 114) = 2.68$, $p = .073$, suggesting that the effect of training format is equal for the age groups. The low p value indicates an interesting trend towards an interaction, however.

For none of the contrasts, the IQ covariate reveals main effects or interactions (p-values > .05). It was therefore discarded from the contrast analyses. For the contrast CP-MWE, there are both effects of age group, $F(1, 76) = 5.94$, $MSE = 2.09$, $p = .017$, and training format, $F(1, 76) = 16.51$, $p < .001$, suggesting that the elderly participants experience higher levels of training load than the young participants, but that both age groups show lower levels of load in the MWE condition than in the CP condition. Regarding the corrected α-level of .017, there is also a trend towards an interaction, $F(1, 76) = 4.87$, $p = .03$, indicating that the elderly show a greater SCL 'decrease' between the conventional and the multimedia condition than the young.

The contrast UWE-MWE, on the other hand, neither reveals main effects nor an interaction (all p's > .2). This means that the multimedia condition does not reduce the level of SCL relative to the unimodal condition. It is interesting, though, that neither the unimodal nor the multimedia condition reveals any significant age difference in SCL.

The CP-UWE contrast, finally, shows a main effect of age group, $F(1, 76) = 11.15$, $MSE = 1.98$, $p < .01$, as well as training format, $F(1, 76) = 10.37$, $p < .01$. This is a replication of Van Gerven *et al.*'s (2002) finding that worked examples impose less cognitive load on the learner than conventional problems. However, unlike Van Gerven *et al.*, no interaction effect was found ($p > .1$).

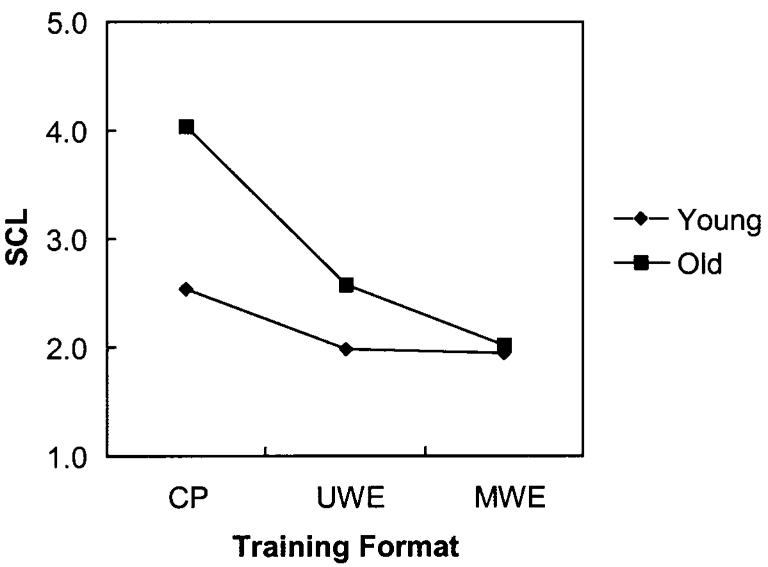

Figure 2. Mean subjective cognitive load (SCL)

Performance

IQ scores have a significant impact on performance, $F(1, 111) = 4.78$, $MSE = .035$, $p < .05$. However, the absence of an IQ × training format interaction, $F(2, 111) = 1.85$, $p > .1$, suggests that IQ has no impact on the effectiveness of the different training formats. The interaction term was therefore discarded from the overall analysis. A main effect of age group was found, $F(1, 113) = 103.39$, $MSE = .035$, $p < .001$, indicating that the young performed better than the elderly. However, there is neither an effect of training format, nor an interaction ($p > .3$).

The contrast CP-MWE reveals a trend towards a main effect of IQ, $F(1, 74) = 4.50$, $MSE = .034$, $p = .037$, as well as a trend towards an IQ × training format interaction, $F(1, 74) = 3.55$, $p = .064$. This indicates that there is a marginal influence of IQ on performance as well as on the effect training format has on performance. Further analysis reveals that, again, there is an effect of age group, $F(1, 74) = 83.55$, $MSE = .034$, $p < .001$. Remarkably, there is trend towards a main effect of training format, $F(1, 74) = 3.79$, $p = .056$, indicating that performance is slightly better in the MWE than in the CP condition (see Table 1). The absence of an age group × training format interaction ($p > .1$) reveals, however, that this trend does not differ between the age groups.

The contrast UWE-MWE reveals neither a main effect nor an interaction for the IQ covariate. Excluding IQ from the analysis only reveals a significant main effect of age group, $F(1, 76) = 66.70$, $MSE = .035$, $p < .001$. No main effect of training format or an age group × training format interaction was found ($p > .2$).

The contrast CP-UWE, finally, reveals a main effect of IQ, $F(1, 74) = 6.33$, $MSE = .035$, $p = .014$. However, no IQ × training format interaction could be detected ($p > .1$). The interaction term was therefore discarded from the contrast analysis. Further analysis only reveals an effect of age group, $F(1, 75) = 78.40$, $MSE = .036$, $p < .001$. There was neither an effect of training format nor an interaction ($p > .3$).

Training efficiency

Figure 3a represents training efficiency as the mean z scores of SCL and performance in a co-ordinate system (following Paas & Van Merriënboer, 1993). Figure 3b depicts the mean values of training efficiency. A trend towards an effect of IQ was detected, $F(1, 111) = 4.66$, $MSE = 1.27$, $p = .033$, but no interaction with training format ($p > .1$). The interaction term was therefore discarded from the overall analysis. Remarkably, further analysis fails to reveal a main effect of age group ($p > .2$). There is an effect of training format, however, $F(2, 113) = 3.98$, $MSE = 1.29$, $p < .05$, but no interaction ($p > .4$). This indicates that for both age groups efficiency is greater for the CLT-based instruction formats than for the conventional format.

IQ appears to have only marginal effects in the CP-MWE contrast analysis: there is a trend towards a main effect, $F(1, 74) = 3.32$, $MSE = 1.24$, $p = .072$, as well as a trend towards an interaction, $F(1, 74) = 4.14$, $p = .046$. The effect of training format, resulting from the overall analysis, can be fully attributed to the CP-MWE contrast, because this is the only contrast for which the 2 × 2 ANCOVA reveals a trend towards a main effect, $F(1, 74) = 4.94$, $p = .029$. That is, the MWE format seems to be more efficient than the CP format. Furthermore, there is no main effect of age group, ($p > .2$) but there is a trend towards an age group × training format interaction, $F(1, 74) = 3.64$, $p = .06$, suggesting that the MWE condition is slightly more advantageous for the elderly than for the young.

The UWE-MWE contrast analysis reveals no significant effects with respect to the IQ covariate, which is therefore excluded from the analysis. Further analysis revealed neither any main effects nor an interaction ($p > .1$). Regarding the results of the CP-MWE contrast analysis, this means that although the multimedia condition does not seem to have a surplus value relative to the unimodal worked examples condition, it apparently does have a surplus value relative to the conventional condition.

With respect to the CP-UWE contrast, there is a clear main effect of IQ, $F(1, 74) = 7.19$, $MSE = 1.29$, $p < .01$, but no interaction with training format ($p > .3$). The interaction term was therefore discarded from the analysis. Further analysis only reveals a trend towards an effect of age group, $F(1, 75) = 3.41$, $MSE = 1.29$, $p = .069$. The absence of an effect of training format in the CP-UWE contrast analysis confirms that the multimedia condition is responsible for the effects found in the overall analysis.

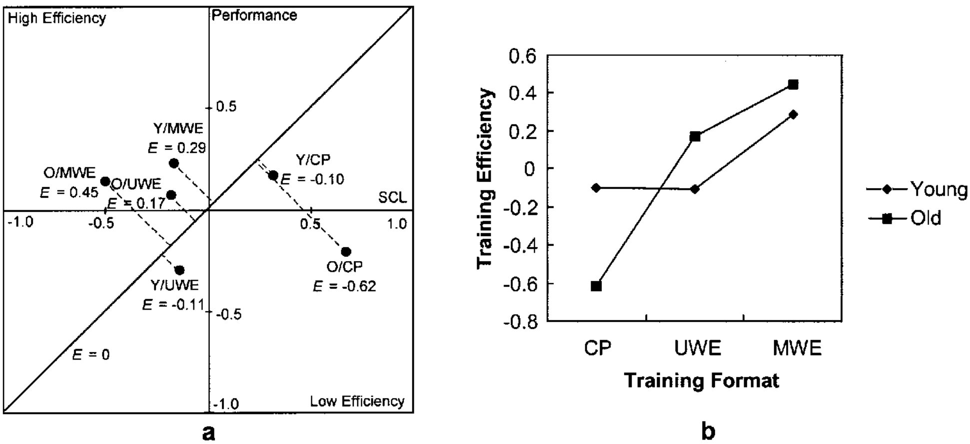

Figure 3. Mean training efficiency

Training time

Mean training time is depicted in Figure 4. IQ has a significant impact on the time spent on training, $F(1, 111) = 9.30$, $MSE = 968.77$, $p < .01$. Furthermore, it interacts with training format, $F(2, 111) = 4.10$, $p < .05$, which means that it influences the effect of training format on training time. There is a main effect of age group, $F(1, 111) = 52.33$, $p < .001$, which denotes that the elderly needed more training time than their young counterparts. Furthermore, there is a main effect of training format, $F(2, 111) = 4.59$, $p < .05$, and an age group × training format interaction, $F(2, 111) = 12.69$, $p < .001$. Regarding Figure 4, this indicates that the training times of the young and elderly participants 'grow' towards each other, which is due to the fixed length of the animations in the MWE condition. That is, the young were compelled to spend a minimum training time in this condition.

The CP-MWE contrast analysis also reveals a significant main effect of IQ, $F(1, 74) = 6.84$, $MSE = 1193.93$, $p = .011$, as well as an interaction, $F(1, 74) = 6.37$, $p = .014$. Furthermore, there are significant main effects of age group, $F(1, 74) = 28.86$, $p < .001$, and training format, $F(1, 74) = 6.86$, $p < .05$, as well as an interaction, $F(1, 74) = 20.60$, $p < .001$, which is in line with the overall analysis.

In the UWE-MWE contrast analysis, IQ shows neither a main effect nor an interaction

($p \geq .1$), and was therefore left out of the analysis. The analysis further shows significant effects of age group, $F(1, 76) = 42.85$, $MSE = 302.32$, $p < .001$, training format, $F(1, 76) = 32.29$, $p < .001$, and an interaction, $F(1, 76) = 22.52$, $p < .001$. Again, this is in line with the overall analysis.

The CP-UWE analysis reveals a significant main effect of IQ, $F(1, 74) = 8.97$, $MSE = 1420.53$, $p < .01$, but no interaction ($p > .1$). Thus, the error term is left out. Further analysis reveals main effects of age group, $F(1, 75) = 54.57$, $MSE = 1450.28$, $p < .001$, and training format, $F(1, 75) = 9.84$, $p < .01$, but no interaction, $F(1, 75) = 1.02$, $p > .3$. Regarding Figure 4, this means that less time is needed in the UWE condition than in the CP condition. Furthermore, it confirms that the lines between the CP and UWE condition run parallel, indicating that there are no differential effects between the age groups.

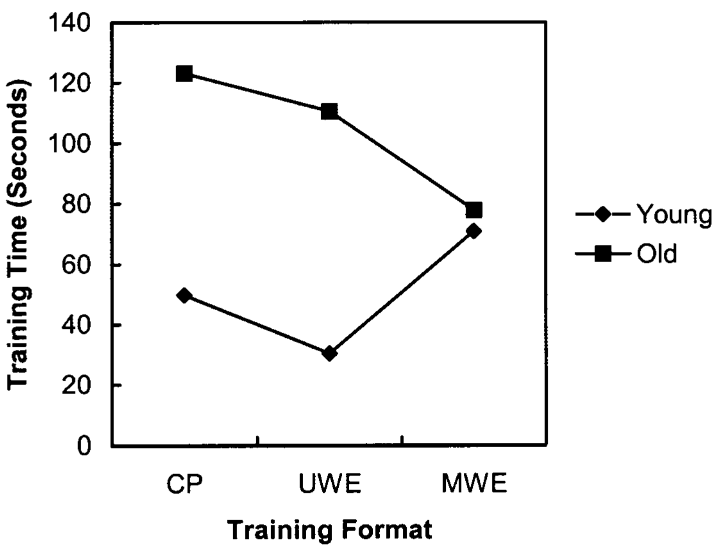

Figure 4. Mean time spent per training problem in seconds

Discussion

The purpose of this study was to investigate the efficiency of training with multimedia-based worked examples relative to training with unimodal worked examples and conventional problems in both young and older adults. Two hypotheses were stated. On the basis of cognitive load theory (Sweller *et al.*, 1998), it was hypothesised that multimedia-based worked examples are more efficient than unimodal worked examples and that unimodal worked examples are more efficient than conventional problems (Hypothesis 1). These effects were expected to be stronger for the elderly than for the young, since the enhancement would be proportionally larger for the former than for the latter (Hypothesis 2; Van Gerven *et al.*, 2000).

Participants studying worked examples of both formats invested significantly less mental effort than participants solving conventional problems, which is in line with Hypothesis 1 and the study by Van Gerven *et al.* (2002). Furthermore, the interaction trend for the comparison between conventional problems and multimedia-based worked examples suggests a disproportional advantage for the elderly participants, which is in line with Hypothesis 2. The analyses of the efficiency scores, which reflected the ratio between invested cognitive load and performance, did not yield this interaction, however. There was a main effect of training format, though. Contrast analyses revealed that this effect was fully attributable to the difference between the conventional and multimedia condition. A remarkable additional finding was that in the contrast analysis of the unimodal and multimedia condition there was no effect of age group, suggesting that the age groups experienced comparable levels of cognitive load. For both age groups, training time was lower in the unimodal worked-examples condition than in the conventional-problems condition. The young participants, however, were compelled to spend more time in the multimedia condition.

The key question to be answered is whether multimedia-based worked examples add to the efficiency of skill training that is aimed at elderly learners. If the definition of efficiency is restricted to the ratio between cognitive load and performance (Paas & Van Merriënboer, 1993), the answer is negative. In a broader perspective, however, there are at least four reasons to take the affirmative view. First, multimedia-based worked examples led to a significantly stronger 'decrease' in perceived cognitive load for the elderly than for the young. Second, the difference in training efficiency turned out to be significant only for the comparison between conventional problems and multimedia-based worked examples, although this applied to both the old and the young participants. Third, there was a trend towards better performance in the multimedia condition relative to the conventional condition, which again applied to both age groups. With respect to training time, finally, there was an interaction between age group and training format for the comparison between the conventional and multimedia condition. That is, the elderly need significantly less time to process multimedia-based worked examples, whereas the young take slightly more training time. In part, this effect is due to the minimum time the young participants had to spend in the multimedia condition. For the greater part, the effect seems to be attributable to the fact that the elderly did not choose to take more time in this condition, in spite of the opportunity to repeat the animations. The significance of training time as a dependent variable is underlined by Pollock, Chandler, and Sweller (2002), who suggest that it can be considered as an additional ingredient of training efficiency.

The ultimate purpose of a training design is not only to make an efficient use of the available processing capacity but also to realise superior performance. The present study provides a first trend in this direction by yielding a slightly better performance in the multimedia than in the conventional condition. Theoretically, further enhancement of performance by means of unimodal or multimedia-based worked examples is attainable. As we have seen, the multimedia condition imposes significantly less cognitive load on the learner than the conventional condition. This offers the opportunity to add so-called *germane* cognitive load (Sweller *et al.*, 1998), which contributes to the construction of cognitive schemata. In future research, this can be tested by increasing the complexity and variability of the training examples (see, e.g., Paas & Van Merriënboer, 1994b), which could eventually increase efficiency by enforcing a substantial performance difference. In its turn, increased efficiency is likely to boost motivation, making lifelong learning not only a necessary but also an enjoyable experience.

Acknowledgements

We are grateful to Saskia Marcano Bakker for recording the narratives and Marijn Bruinink for running part of the experiment. We also wish to thank the staff and students of the Sintermeertencollege in Heerlen, The Netherlands, for their co-operation. Furthermore, we thank an anonymous reviewer for advising us on the statistical analysis. The Netherlands Organisation for Scientific Research (NWO) is gratefully acknowledged for funding this project.

References

Ainsworth, S., Wood, D., & O'Malley, C. (1998). There is more than one way to solve a problem: Evaluating a learning environment that supports the development of children's multiplication skills. *Learning and Instruction*, *8*, 141–157.

Allen, P.A., Madden, D.J., Groth, K.E., & Crozier, L.C. (1992). Impact of age, redundancy, and perceptual noise on visual search. *Journal of Gerontology: Psychological Sciences*, *47*, 69–74.

Authorware 3.5 [Computer software]. (1996). San Francisco, CA: Macromedia.

Baddeley, A.D. (1992). Working memory. *Science*, *255*, 556–559.

Bobis, J., Sweller, J., & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction*, *3*, 1–21.

Bobis, J., Sweller, J., & Cooper, M. (1994). Demands imposed on primary-school students by geometric models. *Contemporary Educational Psychology*, *19*, 108–117.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233–246.

Clark, J.M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210.

Czaja, S.J. (1996). Aging and the acquisition of computer skills. In W.A. Rogers, A.D. Fisk & N. Walker (Eds.), *Aging and skilled performance: Advances in theory and applications* (pp. 201–220). Mahwah, NJ: Erlbaum.

Hasher, L., & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351–371.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, *92*, 126–136.

Luchins, A.S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, *54*, Whole No. 248.

Luteijn, F., & Van der Ploeg, F.A.E. (1983). *Groninger intelligentie test: Handleiding* [Groningen intelligence test: Manual]. Lisse, The Netherlands: Swets & Zeitlinger.

Mayer, R.E., & Anderson, R.B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*, 484–490.

Mayer, R.E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*, 312–320.

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

Mayer, R.E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology*, *88*, 64–73.

McDowd, J.M. (1997). Inhibition in attention and aging. *Journal of Gerontology: Psychological Sciences*, *52*, 265–273.

Moreno, R., & Mayer, R.E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*, 358–368.

Morrell, R.W., & Echt, K.V. (1996). Instructional design for older computer users: The influence of cognitive factors. In W.A. Rogers, A.D. Fisk & N. Walker (Eds.), *Aging and skilled performance: Advances in theory and applications* (pp. 241–265). Mahwah, NJ: Erlbaum.

Morrell, R.W., & Echt, K.V. (1997). Designing written instructions for older adults: Learning to use computers. In A.D. Fisk & W.A. Rogers (Eds.), *Handbook of human factors and the older adult* (pp. 335–361). San Diego, CA: Academic Press.

Morrell, R.W., & Park, D.C. (1993). The effects of age, illustrations, and task variables on the performance of procedural assembly tasks. *Psychology and Aging*, *8*, 389–399.

Mousavi, S.Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*, 319–334.

Myerson, J., Hale, S., Wagstaff, D., Poon, L.W., & Smith, G.A. (1990). The information-loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review*, *97*, 475–487.

Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology*, *77*, 272–284.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*, 429–434.

Paas, F., & Van Merriënboer, J.J.G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, *35*, 737–743.

Paas, F., & Van Merriënboer, J.J.G. (1994a). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, *6*, 351–371.

Paas, F., & Van Merriënboer, J.J.G. (1994b). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, *86*, 122–133.

Paas, F., Camp, G., & Rikers, R.M.J.P. (2001). Instructional compensation for age-related cognitive declines: Effects of goal specificity in maze learning. *Journal of Educational Psychology*, *93*, 181–186.

Paas, F., Van Merriënboer, J.J.G., & Adam, J.J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, *79*, 419–430.

Paas, F., Tuovinen, J.E., Tabbers, H., & Van Gerven, P.W.M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*, 63–71.

Penney, C.G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, *17*, 398–422.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, *12*, 61–86.

Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Salthouse, T.A., & Babcock, R.L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*, 763–776.

ScreenCam [Computer software]. (1997). Cambridge, MA: Lotus.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, *12*, 185–233.

Sweller, J., & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*, 59–89.

Sweller, J., & Levine, M. (1982). Effects of goal specificity on means-ends analysis and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 463–474.

Sweller, J., Van Merriënboer, J.J.G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, *119*, 176–192.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, *3*, 257–287.

Van Gerven, P.W.M., Paas, F., Van Merriënboer, J.J.G., & Schmidt, H.G. (2000). Cognitive load theory and the acquisition of complex cognitive skills in the elderly: Towards an integrative framework. *Educational Gerontology*, *26*, 503–521.

Van Gerven, P.W.M., Paas, F., Van Merriënboer, J.J.G., & Schmidt, H.G. (2002). Cognitive load theory and aging: Effects of worked examples on training efficiency. *Learning and Instruction*, *12*, 87–105.

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, *7*, 1–39.

Wingfield, A., Stine, E.A.L., Lahar, C.J., & Aberdeen, J.S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research*, *14*, 103–107.

Received 23 June, 2001; revised version received 25 August, 2002