Human Factors: The Journal of the Human Factors and Ergonomics Society

http://hfs.sagepub.com/

When Redundant On-Screen Text in Multimedia Technical Instruction Can Interfere With Learning

Slava Kalyuga, Paul Chandler and John Sweller

Human Factors: The Journal of the Human Factors and Ergonomics Society 2004 46: 567
DOI: 10.1518/hfes.46.3.567.50405

When Redundant On-Screen Text in Multimedia Technical Instruction Can Interfere With Learning

Slava Kalyuga, Paul Chandler, and **John Sweller,** University of New South Wales, Sydney, Australia

It is frequently assumed that presenting the same material in written and spoken form benefits learning and understanding. The present work provides a theoretical justification based on cognitive load theory, and empirical evidence based on controlled experiments, that this assumption can be incorrect. From a theoretical perspective, it is suggested that if learners are required to coordinate and simultaneously process redundant material such as written and spoken text, an excessive working memory load is generated. Three experiments involving a group of 25 technical apprentices compared the effects of simultaneously presenting the same written and auditory textual information as opposed to either temporally separating the two modes or eliminating one of the modes. The first two experiments demonstrated that nonconcurrent presentation of auditory and visual explanations of a diagram proved superior, in terms of ratings of mental load and test scores, to a concurrent presentation of the same explanations when instruction time was constrained. The 3rd experiment demonstrated that a concurrent presentation of identical auditory and visual technical text (without the presence of diagrams) was significantly less efficient in comparison with an auditory-only text. Actual or potential applications of this research include the design and evaluation of multimedia instructional systems and audiovisual displays.

INTRODUCTION

Delivering text simultaneously in written and spoken form is common in multimedia presentations. Conference presenters or educators reading their overheads aloud, text on a TV screen with a voice reading the same text, and instructional multimedia including spoken narration with the same written text presented simultaneously on the computer screen are all familiar. Nevertheless, the assumption that simultaneous presentation of written and auditory text is beneficial may be erroneous. In a number of recent studies evaluating the benefits of multimedia instruction (e.g., Beccue, Vila, & Whitley, 2001; Hegarty, Quilici, Narayanan, Holmquist, & Moreno, 1999; Najjar, 1996; Tergan, 1997), redundant information presentations in different modalities (i.e., presenting the same information in written and spoken form) did not bring about the expected positive effects on learning. Establishing relations between different sources of information may be difficult for learners dealing with multiple representations (Van Someren, Reimann, Boshuizen, & de Jong, 1998). We know, from previous work, that some forms of redundancy can interfere with learning (see Mayer, 2001; Sweller, 1999; and Sweller, Van Merrienboer, & Paas, 1998, for reviews). The reasons are embedded deeply in the human cognitive architecture (Sweller, 2003).

Novel information must be handled by working memory and, as is well known, only a few units of information can be processed in working memory at any time (e.g., Baddeley, 1998). Overburdening of working memory may result in decreased effectiveness of information processing. Both a cognitive load theory (Paas, Renkl, & Sweller, 2003; Sweller, 1999; Sweller et al., 1998) and a cognitive theory of multimedia

Address correspondence to Slava Kalyuga, Educational Testing Centre, University of New South Wales, 12-22 Rothschild Ave., Rosebery, NSW 2018 Australia; s.kalyuga@etc.unsw.edu.au. **HUMAN FACTORS,** Vol. 46, No. 3, Fall 2004, pp. 567–581.

learning (Mayer, 2001; Mayer & Moreno, 2003) have been developed to explore the instructional consequences of this fundamental feature of human memory. Working memory may be overburdened if instruction involves excessive elements of novel information processed simultaneously. However, there may be no limitation in the number of familiar, well-learned elements that can be processed in working memory (Ericsson & Kintsch, 1995; Sweller, 2003). As a consequence of this distinction between familiar and unfamiliar material, novices who deal with novel information and experts who deal with familiar information process that information in different ways.

Knowledge is held in long-term memory in the form of hierarchically organized schemas, allowing experts to treat many elements of information as a single element, thus reducing demands on working memory. Because people have a limited information processing capacity, appropriate allocation of cognitive resources is important to efficient learning, especially for relative novices in a domain. In situations where a significant share of mental resources is assigned to activities not directly related to schema acquisition, learning may be inhibited.

Dual-processing models of memory consider capacities to be distributed over separate auditory and visual channels (Baddeley, 1998; Penney, 1989; Schneider & Detweiler, 1987). For example, in Baddeley's (1998) model, the phonological loop processes auditory information (verbal or written material in an auditory form), whereas the visual-spatial sketch pad deals with visual information such as diagrams and pictures. Paivio's (1990) dual coding theory also suggests that information can be encoded, stored, and retrieved from two fundamentally distinct systems, one suited to verbal information, the other to images. Penney (1989) proposed a model of working memory (the "separate stream hypothesis") in which the processing of auditory and visually presented verbal items is carried out independently by auditory and visual processors in working memory.

Dual-mode presentations may effectively expand working memory capacity if one part of the instruction (e.g., textual explanations) is presented in auditory form and the other (e.g., a diagram) in visual form, increasing the amount of information that can be processed without cognitive overload. Mayer and Moreno (1998) and Moreno and Mayer (1999; see Mayer, 2001, for a recent review), Mousavi, Low, and Sweller (1995), and Tindall-Ford, Chandler, and Sweller (1997) demonstrated the superiority of audiovisual instructions. Learners integrated words and pictures more easily when the words were presented auditorily rather than visually because using auditory and visual processors in working memory effectively eliminated cognitive overload of the visual channel. Mayer and Anderson (1992) and Mayer and Sims (1994) found that concurrent presentations of pictorial and verbal information were superior to sequential presentations (the split-attention or contiguity effect). They demonstrated that dual-modality instructions (animations accompanied by audio text) were a better instructional format only when the audio and visual components were presented simultaneously rather than sequentially. However, all the instructional components in those studies were nonredundant. An essential audio text accompanied essential visual graphic information. A possible reversal of the split-attention or contiguity effect, with audiovisual presentation interfering with learning under conditions of verbal redundancy, needs to be investigated.

Generally, redundancy effects may occur when learners are required to integrate several sources of information, including their own schematic knowledge, that have an identical information content but a different surface structure. If a learner can successfully handle a situation or task using one source of information, presenting her or him with other sources that simply redescribe the same subject may cause a cognitive overload. Attending to redundant information consumes cognitive resources that become unavailable for learners to process essential information. Eliminating redundant information frees these resources for learning (Chandler & Sweller, 1991; Sweller & Chandler, 1994). Some material may also become redundant because of an increased level of expertise in a domain (Kalyuga, Chandler, & Sweller, 1998).

Kalyuga, Chandler, and Sweller (1999, 2000) observed conditions where the addition of concurrent audio explanations to visual instructions had negative rather than positive or neutral effects. The measures of cognitive load used in those studies suggested that those conditions occurred when processing an auditory supplement was likely to impose an excessive load on working memory. We assumed that the need to attend to, coordinate, and process both modes of text simultaneously, and to relate them to other graphic information, consumed additional cognitive resources and thereby overloaded working memory capacity and hindered learning. Similar results were obtained by Moreno and Mayer (2002) and by Mayer, Heiser, and Lonn (2001), who came to the conclusion that concurrent on-screen text, animations, and narrations overloaded the visual channel because of competition between the animation and visual text for cognitive resources.

The aforementioned results challenge the commonsense view that presentations of the same written and auditory text may be beneficial under some circumstances simply because the material is presented twice. Any such advantage may disappear during simultaneous presentation because of working memory overload but may be hypothesized to reappear if simultaneous presentations are not used. In many cases, a redundant source of information may be separated in time and not need to be processed together with a primary source (e.g., revision of previously learned material). Processing redundant material in this situation should not increase working memory load and may well be useful for learning. However, if the same two sources of information are organized to be processed simultaneously (e.g., listening to and reading the same text), concurrent processing of modules of identical information might exceed working memory capacity and thus decrease the effectiveness of learning. This type of instructional redundancy was of prime interest in the experiments reported here.

It was hypothesized that if audio and visual contents are presented serially rather than simultaneously, the two versions of the text will not require concurrent processing and coordination in working memory. In fact, the two versions of text presented serially may complement and consolidate learning of instructional material. Cognitive resources in this case might not be diverted to establishing relations between corresponding visual and auditory elements. The experiments presented in this paper were designed to investigate whether nonconcurrent presentation of audio and visual modes of the same textual information would improve learning in comparison with concurrent audiovisual presentation of textual explanations.

The experiments were conducted in realistic environments within the training facilities of two large Australian manufacturing companies. The experimental materials for each experiment were parts of different training modules for the same group of apprentices at several separate stages throughout their 1st year of training. Because of this variation in content, each study was considered to provide a new content that was not familiar to the trainees from their previous learning experience, allowing the same participants to be used in the experiments. Experiment 1 compared concurrent (auditory and visual) and sequential (auditory followed by visual) presentations of textual explanations of a diagram and had no limitation on available instruction time. Experiment 2 compared similar instructional formats (using different materials) under constrained instruction time conditions. To eliminate a possible influence of perceptual difficulties when processing visual written explanations of a visual diagram, in the third experiment we compared simultaneous presentation of audio and visual text with an audio-alone condition without a diagram and predicted again that the elimination of the redundant visual mode would facilitate learning.

EXPERIMENT 1

The training materials for Experiment 1 were in the area of light fabrication and, in particular, the reading of cutting speed charts. Specifically, trainees were required to learn how to use a Cutting Speed Nomogram, which is used to determine the appropriate number of revolutions per minute (R.P.M.) to run a drill of a given diameter (in millimeters) at a given cutting speed (in meters/min). The recommended cutting speed ranges for different materials were available to students in a separate table. A section of a computer screen presentation for instruction on the cutting speed nomogram is presented in Figure 1.

Assume that the diagram of Figure 1 with text explaining how to use the nomogram to

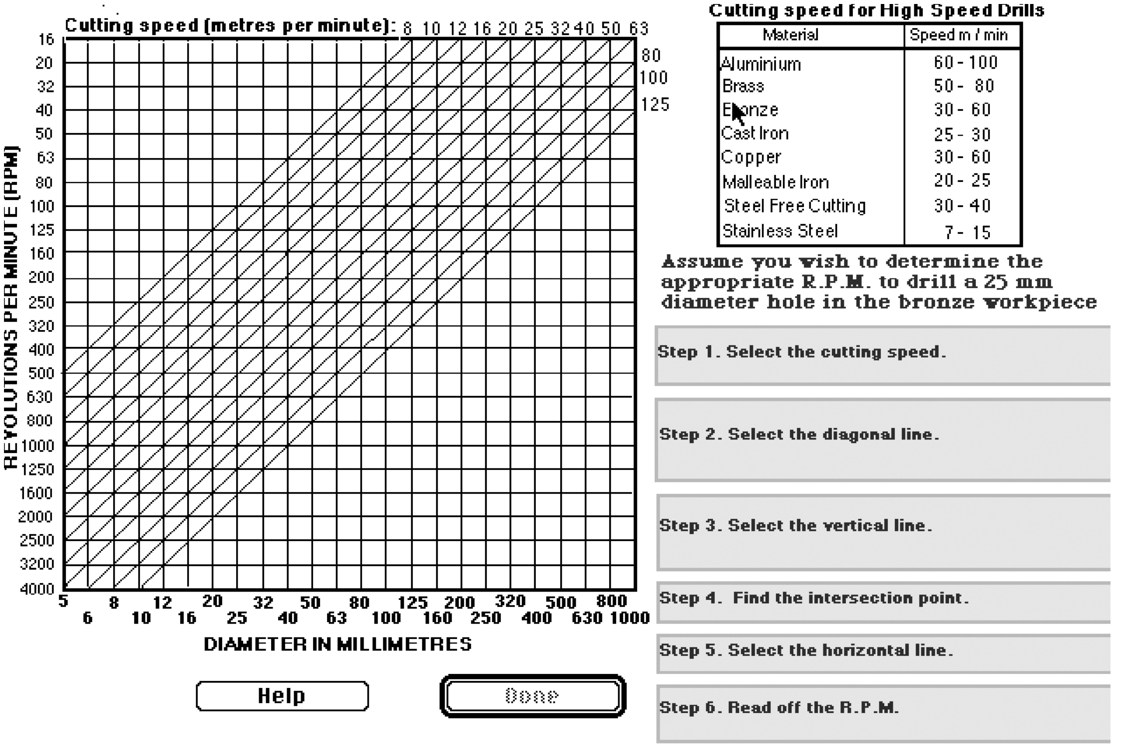

Figure 1. A section of computer screen presentation for instruction on the cutting speed nomogram. Adapted from Kalyuga, Chandler, and Sweller (2000). Copyright © 2000 by the American Psychological Association.

determine R.P.M. for a specified diameter and particular material is presented to relatively inexperienced learners who understand the basic concepts used in the instruction (cutting speed, R.P.M., etc.) but have never used the nomograms before. If the visual text is used simultaneously with an auditory narration of the same text, working memory capacity might become overloaded because of redundancy. When reading and listening to the same verbal material simultaneously, learners must establish connections between corresponding elements of visual and auditory components of the text. The process of coordination of the two sensory modes may unnecessarily consume additional cognitive resources, as compared with those used in a format with a diagram plus auditory text alone. Additional resources required for coordination will be unavailable for learning.

One way to reduce the possible negative instructional consequences of such redundant information might be to present the visual text only after the auditory explanation has been fully articulated, not simultaneously with an auditory narration of the text. In this situation, visual and auditory explanations need not be mentally integrated in working memory and thus do not compete for working memory resources. Working memory capacity is not wasted on establishing connections between corresponding elements of visual and auditory components and a precise coordination of the two sensory modes. Working memory resources, otherwise used for such coordination of visual and auditory text, will be available for learning.

Thus, from the point of view of cognitive load theory, a nonconcurrent duplication of text-based information using different modes of presentation might not increase the risk of overloading working memory capacity and should not have the potential negative learning consequences of concurrent auditory and visual text. In this experiment we compared two instructional formats: a diagram with audio text and concurrent visual text, and a diagram with audio text and delayed (nonconcurrent) visual text. In accord with cognitive load theory and the previous discussion, it was expected that trainees would benefit more from the nonconcurrent presentation format than from the concurrent one.

Method

Participants. Twenty-five trade apprentices 16 to 19 years of age participated in this experiment. All participants had completed at least Year 10 of high school and had completed about 2 months of their 1st-year trade course. During their regular training courses, all participants were introduced to the technical terminology, cutting principles, and equipment necessary to understand the instructional materials. None of the participants had previous exposure to cutting speed nomograms used for calculating R.P.M. (tables had been used for this purpose during practical work).

Materials and procedure. Participants were randomly allocated to two groups corresponding to the two instructional formats. Thirteen learners were allocated to the concurrent text group, and 12 learners were allocated to nonconcurrent text group. All instructions and training for the study were delivered via an Apple Power Macintosh computer, for which the first author had designed all the computer-based training packages using Authorware Professional. All participants were tested individually.

The concurrent text format (see Figure 1) contained a cutting speed nomogram and the headings of the sequential steps involved in using this nomogram (e.g., "Step 1. Select the cutting speed. Step 2. Select the diagonal line."). A textual problem statement ("Assume you wish to determine the appropriate R.P.M. to drill a 25 mm hole in the bronze workpiece") was presented visually next to the nomogram. When a trainee clicked on a particular step, auditory narration of an explanation of this step was delivered through headphones and visual text was displayed next to the step number. The auditory-based text was identical to the visual text information, and the two were delivered simultaneously. Visual highlights of the appropriate elements of the nomogram (the material name, lines, intersection points, etc.) and some animations of the diagrammatic information were presented simultaneously with corresponding explanations. For example, if the learner clicked on "Step 4. Find the intersection point," the sentence "Follow the diagonal line until it intersects with the vertical line" was displayed and articulated and an arrow would move along a highlighted diagonal line from the top toward the highlighted intersection point. Learners were required to attend to all the procedural steps at least once. (In fact, computer records indicated that all learners in both groups attended to each step only once.)

The nonconcurrent text format was visually identical to the concurrent text format except that spoken explanations for each step were presented first, followed immediately by the identical written explanations. Written explanations for each step were presented after the corresponding spoken explanations for this step were fully articulated. Thus the written text followed the spoken text in a step-by-step manner. In both conditions, on-screen text for each step remained on the screen as long as a learner wished (until he or she clicked for the next step to be explained), thus equalizing conditions in terms of the exposure to the textual explanations. To allow learners to choose their own instruction pace, time was not controlled in this experiment.

Experimental training. Trainees studied the cutting speed nomogram instructional material (see Figure 1) in their respective experimental groups. All computer-based presentations were self-paced, and there was no time limit for this instruction phase. The time each learner spent on studying instructions was recorded electronically on the computer.

Subjective ratings. After the participants studied the instructions, subjective ratings of task difficulty (see Paas & Van Merrienboer, 1993, 1994) were collected from all participants electronically on the computer. A 7-point scale was used for the question "How easy or difficult was this nomogram to understand? Click your answer." The participants selected one of the seven options (*extremely easy, very easy, easy, neither easy nor difficult, difficult, very difficult,* and *extremely difficult*). Thus a mental load rating ranging from 1 (*extremely easy*) to 7 (*extremely difficult*) was collected for each participant. Such rating scales are increasingly being used as an effective and valid measure of the subjective mental load related to a particular learning task (see Kalyuga et al., 1998, 1999, 2000; Mayer

& Chandler, 2001; Paas, 1992; Paas & Van Merrienboer, 1993, 1994; Tindall-Ford et al., 1997). Although this self-report technique cannot be regarded as a unique measure of working memory load, it is the most suitable one in conditions of scheduled sessions in industrial training centers because of its minimal interference with training procedures. (For discussions of techniques for measuring cognitive load, see Brunken, Plass, & Leutner, 2003; Paas, Tuovinen, Tabbers, & Van Gerven, 2003.)

Performance test. Ten multiple-choice questions (with four alternatives for each question) followed subjective ratings of task difficulty. Participants had to find an unknown variable from given variables using the cutting speed nomogram. Answering the questions required applying various combinations and sequences of procedural steps, most of which were different from those presented during the experimental training and so were transfer questions. For example, the question "Assume a drilling machine with the range of R.P.M. between 16 and 800, and a drill of 8 mm diameter. Which materials from the above table can be drilled using this machine?" with the choices of "free cutting steel," "stainless steel," "copper," and "cast iron" involved (a) selecting a horizontal line (800 R.P.M.) as the lower border of the R.P.M. range (the upper border in this case was the upper edge of the nomogram); (b) selecting a vertical line (8 mm); (c) locating the intersection points of the vertical line with diagonal lines within the given range of R.P.M.; (d) determining a range of cutting speed values corresponding to the diagonal lines running through the located intersection points (from 8 to 20 m/min); and, finally, (e) selecting from the table the materials that included values from that range in their cutting speed range (in this case, only stainless steel). Up to 1.5 min were allowed for each of the 10 questions. The responses to each question were electronically recorded and judged as either correct or incorrect, providing a score out of 10 for each participant.

Results and Discussion

The independent variable was the instructional format (concurrent or nonconcurrent text). The dependent variables were instruction time (the time each learner spent on studying the instructions during the experimental training), subjective ratings of mental load, test performance scores on the multiple-choice items, and instructional efficiency measures. Means and standard deviations are displayed in Table 1.

Instructional efficiency measures were calculated using Paas and Van Merrienboer's (1993, 1994) procedure, which combines subjective measures of cognitive load with measures of

TABLE 1: Mean Instruction Times, Ratings, Scores, and Efficiencies by Experimental Condition for Experiment 1

	Experimental Condition	
Variable	Concurrent Text	Nonconcurrent Text
Instruction time (s)		
M	133.6	156.1
SD	39.9	66.9
Rating scale		
M	2.7	1.6
SD	1.6	0.8
Questions scores		
M	5.8	6.8
SD	1.8	2.2
Instructional efficiency		
M	–0.45	0.49
SD	1.32	0.94

Note. Maximum questions score = 10. Ratings were made on 7-point scales (1 = *extremely easy*, 7 = *extremely difficult*). Actual range for instructional efficiency was from –4.13 to 1.91 (possible range is $-\infty$, $+\infty$).

test performance in order to determine the relative efficiency of instruction. It assumes that instructional presentations are efficient if they produce better performance results with less cognitive load. Efficiency values were calculated by converting cognitive load and performance measures into z scores (standardizing those measures across conditions) and combining z scores using the formula $E = (P - R)/\sqrt{2}$, in which E = efficiency, P = performance z score, and R = rating scale z score (with the $\sqrt{2}$ in the denominator being used to make the graphical interpretation of the formula more straightforward; see Paas & Van Merrienboer, 1993, for details). If performance and rating z scores are equal ($P = R$), efficiency is zero ($E = 0$). If the performance z score is higher than the rating z score ($P > R$), instructional efficiency is positive ($E > 0$), but if the performance z score is lower than the rating z score ($P < R$), instructional efficiency is negative ($E < 0$).

The use of the efficiency measure allowed us to be more confident about subjective ratings as measures of cognitive load. If a learner indicated a low mental effort but performed well on the test, it is more probable that those ratings reflected cognitive load rather than irrelevant subjective feelings concerning the materials. In addition, efficiency measures, by combining performance and subjective ratings measures, provided us with a single, instructionally relevant measure of the effects of instruction.

Multivariate analysis of variance (MANOVA) with instruction time, subjective ratings of mental load, and test performance scores as multiple dependent variables (efficiency was not included in this analysis because it is a derivative variable calculated from test performance and subjective rating scores) indicated no significant effect, $F(3, 21) = 1.98$ (Hotelling's trace value of .30). Despite the nonsignificant difference, it might be noted that there was a significant difference for subjective ratings of mental load, $t(23) = 2.15$, $p < .05$ on a one-tailed test (hypotheses were directional). When performance and ratings of mental load were combined into a single, instructional efficiency measure, results also indicated a significant difference, $t(23) = 2.01$, $p < .05$. The instructional format based on a nonconcurrent presentation of audio and visual text was more efficient than the concurrent presentation format. It should be noted that Cohen's f effect size indices were 0.86 for subjective rating data and 0.80 for instructional efficiency (both indicating large size effects).

Although some of the expected results were obtained in this experiment, they were generally weaker than expected. No statistically significant effect was found for the multiple-choice items (a main performance indicator). The effect may have been weakened because instruction time was not limited in this experiment and the trainees themselves determined the pace of the instruction. In both the concurrent and nonconcurrent conditions, the learners could study visual explanations as long as they needed after the audio narration had ended. The extended exposure to visual instructions could compensate for cognitive overload during the preceding learning episode. As a result, redundancy may have made trainees feel overloaded without affecting their performance, at least on tasks that were not completely new for them.

The trainees' control over the pace of instruction could be an important factor influencing our results. For example, Tabbers, Martens, and Van Merrienboer (2001) showed that replacing on-screen text with audio narration (the modality effect) was effective only when the pace of instruction was set by the time of the narration and students had no control over the pacing. Mayer and Chandler (2001) also demonstrated that increasing learner control over the pacing of instruction (by adding some user interaction to multimedia animation) significantly improved subsequent learner performance. A second experiment was designed to further test our hypothesis that nonconcurrent text presentation would be superior to concurrent presentation by excluding the influence of the pacing factor using constrained instruction time conditions.

EXPERIMENT 2

The second experiment also used training materials in the area of fabrication, but in this study the specific instructional materials were in the domain of soldering and interpreting soldering diagrams. Specifically, for this experiment we used a fusion soldering diagram demonstrating some characteristics of solder. Figure 2 depicts a section of a computer screen for instruction on the fusion diagram used to represent states

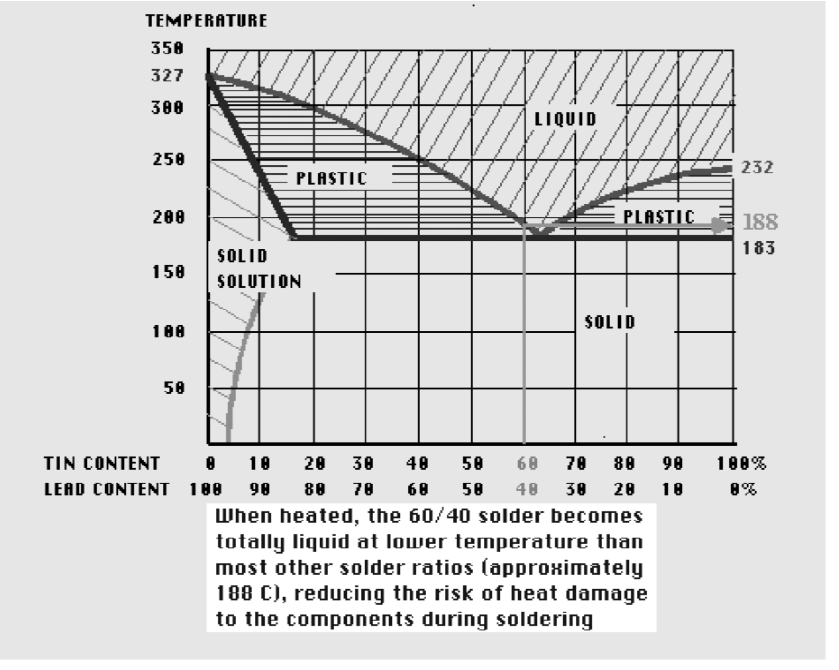

Figure 2. A section of computer screen presentation for instruction on the fusion diagram. Adapted from Kalyuga, Chandler, and Sweller (1999). Copyright ©1999 by John Wiley & Sons, Ltd.

(solid, plastic, or liquid) of tin-and-lead solder, depending on its content (tin/lead ratio) and temperature. It is used to determine the most appropriate solder to use for specific needs and conditions.

The instructional material of Figure 2 (explaining a feature of 60/40 solder) includes many complex interacting elements of information (e.g., 60/40 ratio, temperature value, liquid state, plastic state, heating that corresponds to an upward move along the ratio line, the borders of the plastic state), so we expected a heavy cognitive load to be imposed. If the visual text is used concurrently with an oral narration of the text, all the interacting elements of the diagram and the visual and auditory explanations must be mentally integrated, imposing a heavy cognitive load. If instruction time is limited by a preset presentation time, we can expect that the effect of cognitive overload could not be offset by the longer study time. As indicated previously, that cognitive load could be reduced by a temporal separation of the written and auditory components of the text. We expected that in these conditions a nonconcurrent version of textual information would be superior to a concurrent presentation.

Method

Participants. Twenty-one trade apprentices participated in this experiment. Most of them had participated in Experiment 1 a month prior. (There were several absentees from the previous group and several newcomers.) Most apprentices had some very basic practical experience with soldering, but none had any previous experience reading and interpreting fusion diagrams.

Materials and procedure. Participants were randomly allocated to two groups corresponding to the two instructional formats: 10 learners in the concurrent text group, and 11 learners in the nonconcurrent text group. All instructions and training for the study were delivered via an Apple Power Macintosh computer, for which the first author had designed the computer-based training packages using Authorware Professional. All participants were tested individually.

Both formats contained identical, sequentially introduced, animated components of the fusion diagram (axes, curves, different areas of the diagram, etc.) with auditory explanations (presented via headphones) of newly appearing elements. Descriptions of major features of the 60/40 and eutectic solders (the two most widely used solders) followed. With the concurrent text format, the same portions of explanations were simultaneously presented in a visual format. The nonconcurrent text format differed from the concurrent text format only in that the sections of visual text were presented immediately after the corresponding portions of auditory explanations were fully articulated, rather than simultaneously. The visual text in this format was presented in a step-by-step manner for the same amount of time as that required to articulate the corresponding portions of auditory explanations. In contrast to Experiment 1, in this experiment the instruction time for each section of presentation was preset and identical for both groups (in the concurrent text group, the diagram remained on the screen before the next instructional section for an amount of time equal to that of the visual text exposure in the nonconcurrent group).

Experimental training. Trainees studied the fusion diagram instructions in their respective experimental groups.

Subjective ratings. After participants studied the instructions, subjective ratings of task difficulty were collected from all participants using the same techniques as those used in Experiment 1. Instead of the 7-point scale of Experiment 1, however, a 9-point scale was used. The options were *extremely easy, very easy, moderately easy, slightly easy, neither easy nor difficult, slightly difficult, moderately difficult, very difficult,* and *extremely difficult*). A mental load rating ranging from 1 (*extremely easy*) to 9 (*extremely difficult*) was therefore collected for each participant. A 9-point scale was adopted in this study in order to reduce the extent to which participants used ratings at the extreme ends of the scale and to provide more options.

Performance test. A series of 10 multiple-choice questions followed. The fusion diagram was presented on the screen. The first five questions were directly concerned with the features of solders that had been described previously in the instructional materials (e.g., "What do you think is the most valuable characteristic of any good solder?" "Why is the eutectic solder most appropriate for using in difficult production conditions?" "What does this temperature [1830] mean for the eutectic solder?"). The last five questions required learners to apply their knowledge of the fusion diagram to calculate a specific numerical value or compare several different values and so were transfer questions (e.g., "What is the lowest temperature at which the 20/80 solder becomes plastic?" "Which solder becomes liquid at a lower temperature: 30/70 or 70/30?" "Which solder from those listed below becomes liquid at a lower temperature?"). Up to six numerical or verbal alternatives were provided for each multiple-choice question. Up to 60 s were allowed for each of the 10 questions, and a clock was provided on the screen to indicate the time remaining. The responses to each multiple-choice item were electronically recorded and judged as either correct or incorrect, providing a score out of 10 for each participant.

Results and Discussion

The independent variable was the instructional format (concurrent or nonconcurrent text). The dependent variables were subjective ratings of mental effort, test performance scores on multiple-choice items, and instructional efficiency measures. As in Experiment 1, instructional efficiency measures were calculated using Paas and Van Merrienboer's (1993, 1994) procedure. Means and standard deviations are displayed in Table 2.

MANOVA with subjective ratings of mental load and test performance scores as multiple dependent variables indicated a significant effect, $F(2, 18) = 3.58$, $p < .05$ (Hotelling's trace value of .39). Because all hypotheses were directional, one-tailed t tests of between-subjects effects were performed on the data. There were significant differences between groups for subjective ratings of mental load, $t(19) = 2.27$, $p < .05$, and for multiple-choice items, $t(19) = 1.86$, $p < .05$. Combining these two measures into a single, instructional efficiency measure indicated a significant effect, $t(19) = 2.70$, $p < .01$. The instructional format based on nonconcurrent presentation of audio and visual text was significantly more efficient than the concurrent presentation format. Cohen's f effect size indices

TABLE 2: Mean Ratings, Scores, and Efficiencies by Experimental Condition for Experiment 2

Variable	Experimental Condition: Concurrent Text	Nonconcurrent Text
Rating scale		
M	4.0	2.6
SD	1.7	1.2
Questions scores		
M	5.6	6.9
SD	1.6	1.6
Instructional efficiency		
M	–0.62	0.56
SD	1.09	0.91

Note. Maximum questions score = 10. Ratings were made on 9-point scales (1 = *extremely easy*, 9 = *extremely difficult*). Actual range for instructional efficiency was from –2.16 to 2.11 (possible range is $-\infty$, $+\infty$).

were 0.99 for subjective rating data, 0.81 for the multiple-choice items, and 1.18 for instructional efficiency. All of these values indicate large size effects.

In summary, the concurrent text group performed significantly worse than the nonconcurrent text group on the multiple-choice test. At the same time, the concurrent text format was rated higher in subjective mental load than was the nonconcurrent text format. Thus a redundancy effect was obtained under conditions that required learners in the concurrent text group, for a preset time, to read the text and simultaneously listen to auditory explanations on how to use a diagram for solving a specific problem. Attending to redundant explanations in this case imposed an additional cognitive load, thus decreasing performance and the efficiency of the concurrent instructional presentation.

Delayed presentation of visual text in the nonconcurrent format, which does not require additional working memory load, may also effectively transform this presentation into a form of revision of previously learned auditory presented material. The revision may enhance the advantages of nonconcurrent presentation because of factors not related to working memory load. It is important to investigate whether the redundancy effect would remain in place even when the positive influence of a revised component of instruction is eliminated. Also, in Experiments 1 and 2, diagrams were used as essential parts of instruction. Diagrams were displayed continuously in both experimental groups (concurrent and nonconcurrent formats) and could represent a factor potentially influencing the results (e.g., inspecting the diagram could distract learners from reading the text, thus imposing a perceptual load rather than a cognitive load). Experiment 3 was designed to investigate whether a redundancy effect would be obtained using textual-only materials with no revision involved.

EXPERIMENT 3

In this experiment, a visual text used concurrently with an auditory narration of the same text was compared with auditory-only text. It was hypothesized that the interacting elements of the visual and auditory explanations in the concurrent presentation format had to be mentally integrated, thereby imposing a heavy cognitive load because of redundancy. Cognitive load should be reduced by complete elimination of the redundant written text. We expected that a nonredundant auditory-only textual format should be superior to a concurrent presentation of the same auditory and visual text.

Method

Participants. Twenty-one trade apprentices participated in this experiment. Almost all of them had participated in the previous experiment several months prior. During their regular training courses, all participants were introduced to the technical terminology and equipment necessary to understand the instructional materials, but none of them had any previous exposure to the instructional materials used in the experiment.

Materials and procedure. Participants were randomly allocated to two groups corresponding to the two instructional formats: 11 learners in the concurrent text group and 10 learners in the auditory-only text group. All instructions and training for the study were delivered via an Apple Power Macintosh computer. All participants were tested individually. Each participant spent about 35 min working at the computer.

The training materials used in this experiment were in the area of basic mechanical engineering. Four different sections of text (around 300–400 words each) that did not require pictorial information were prepared. Two sections (about machine frames and tool wear), both of which were adopted from mechanical engineering textbooks that were not used in the training centers, were in a domain that was familiar to participants from their regular training courses and practical work. The instructions contained additional information that had not been emphasized during previous training. Another two sections (about prestressed concrete and underwater welding), which were both adopted from a Reader's Digest popular science and technology book, were not directly related to previous training courses but did not require any specialized prerequisite knowledge.

Both formats contained identical oral narrations of the same sections of text presented via headphones. With the concurrent text format, the same explanations were simultaneously presented in visual form on the screen. Thus, similar to Experiment 2, in this experiment the instruction time was preset and identical for both groups. Identical procedures were used for each of the four sections of text.

Trainees studied a section of text in their respective experimental groups. After studying the instructions, subjective ratings of mental load were collected electronically from all participants using the same techniques as those used in Experiment 2.

Performance test. A series of eight multiple-choice questions followed each section of text, making total of 32 questions in the test. The questions were directly concerned with the previously described factors (e.g., "What is a major requirement for the design of a machine frame?" "If the price of mild steel is about half that of cast iron, why are mild steel frames often more expensive?" "Why may producing a single cast iron frame not be cost effective?"). Four alternatives were provided for each multiple-choice question. Up to 45 s were allowed for each of the questions, and a clock was provided on the screen to indicate the remaining time. The responses to each multiple-choice item were electronically recorded and judged as either correct or incorrect, providing a score out of 32 for each participant.

Results and Discussion

The independent variable was the instructional format (concurrent or auditory-only text). The dependent variables were test performance scores on multiple-choice items, subjective ratings of mental effort averaged over the four sections of text, and instructional efficiency measures. Means and standard deviations are displayed in Table 3.

MANOVA with subjective ratings of mental load and test performance scores as multiple dependent variables indicated a marginally significant effect, $F(2, 18) = 3.06$, $p = .07$ (Hotelling's trace value of .34). No statistically significant differences between groups were obtained for subjective ratings of mental load, $t(19) = 1.04$. There were significant differences for multiple-choice items, $t(19) = 2.63$, $p < .05$. Combining both measures into a single instructional efficiency measure indicated a significant effect, $t(19) = 2.33$, $p < .05$. The instructional format based on an auditory-only presentation of text was significantly more efficient than the concurrent audio and visual presentation format. Cohen's f effect size indices were 1.15 for the multiple-choice items and 1.02 for instructional efficiency (both indicating large size effects).

These results indicate that a redundancy effect was obtained under conditions that required learners in the concurrent text group to read and listen to the same text simultaneously for a limited time and that required learners in the auditory-only text group to listen to identical auditory explanations for an equal amount of time without reading them. Having the same

TABLE 3: Mean Ratings, Scores, and Efficiencies by Experimental Condition for Experiment 3

Variable	Experimental Condition: Concurrent Text	Experimental Condition: Auditory-only Text
Rating scale		
M	3.8	3.4
SD	1.1	0.7
Questions scores		
M	21.6	25.2
SD	3.3	3.0
Instructional efficiency		
M	–0.49	0.54
SD	1.15	0.84

Note. Maximum questions score = 32. Ratings were made on 9-point scales (1 = *extremely easy*, 9 = *extremely difficult*). Actual range for instructional efficiency was from –3.21 to 2.15 (possible range is $-\infty$, $+\infty$).

information presented in two modes simultaneously is less effective than when it is presented in one mode alone under conditions in which the pacing of instruction is controlled by the system.

GENERAL DISCUSSION

The results of the first two experiments indicate that simultaneous presentation of identical written and auditory material has deleterious effects on learning, as compared with sequential presentation modes, when instruction time is constrained. It was hypothesized that the reason for this effect was that simultaneous presentations overloaded working memory, resulting in neither mode being processed adequately. In contrast, sequential presentations permitted both modes to be handled without a strain on working memory, with the second presentation being used to bolster the positive effects of the first presentation. Evidence for a cognitive load explanation of the results came from subjective ratings, which consistently demonstrated that concurrent presentations were seen as higher in perceived mental effort than were nonconcurrent presentations. The third experiment demonstrated a similar effect, even when the positive influence of the repeated presentation of the text was excluded.

Given the small sample sizes involved in the reported experiments, the power of the tests was quite limited. Nevertheless, the fact that the effects could be obtained and, more importantly, replicated using small sample sizes indicates the strength and stability of the findings.

These results need to be integrated with those of Mayer and Anderson (1991, 1992) and Mayer and Sims (1994), who found that dual-modality instructions were superior only when the audio and visual components were presented simultaneously rather than sequentially. On the surface, these results appear to contradict those reported in the present paper. In fact, there is no contradiction, provided one clearly delineates the split-attention and redundancy effects.

Contrasts Between the Split-Attention and Redundancy Effects

The split-attention effect occurs when two or more sources of information must be integrated before they become intelligible (e.g., a geometric diagram and its associated statements). Learners must split their attention between the sources of information and mentally integrate them. A single source is difficult or impossible to understand in isolation. As a consequence, if two such sources of information are presented with a temporal separation, the working memory load imposed by having to hold one source of information while waiting for and then processing and integrating the second source with the first source may be overwhelming. These were the conditions that applied to the work of Mayer and Anderson (1991, 1992) and Mayer and Sims (1994). Under such conditions, concurrent presentation is superior.

Redundant material differs from material that can lead to the split-attention effect in that both sources of information are intelligible in isolation and do not have to be mentally integrated to be intelligible. One source merely redescribes the other source in a different form or mode. Nothing is gained by presenting both sources simultaneously; indeed, working memory load is increased, rather than decreased, by concurrent presentation because both sources of information are likely to be unnecessarily attended to and integrated. Mental integration increases extraneous cognitive load. Such materials were used in the current experiments.

The distinction between disparate sources of information that are or are not intelligible in isolation can easily be forgotten. It is a critical distinction that leads to the distinct phenomena of the split-attention and redundancy effects. Which of these two effects is likely to occur should determine instructional design. Split-attention effects can be ameliorated by integrating materials both physically and temporally. Redundancy effects can be ameliorated by separating or even eliminating redundant sources of information. Both effects are attributable to working memory limitations and can be explained by cognitive load theory. (Sweller, 1999, and Sweller et al., 1998, provide reviews of both effects.)

The level of learner expertise in a domain might be another factor influencing relations between the split-attention and redundancy effects. The contiguity effects seen by Mayer, Steinhoff, Bower, and Mars (1995) were usually seen only for novice learners (Mayer, 2001). Textual explanations of a diagram, which are essential for a novice, may be redundant for someone with more domain-specific knowledge (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga et al., 1998). For example, auditory explanations of the diagrams in our experiments could also become redundant when presented to more experienced learners under system-paced conditions. Kalyuga et al. (2000) demonstrated that if experienced learners attend to such redundant auditory explanations, learning might be inhibited in comparison with a diagram-alone condition.

The Size of Text Segments

Thus, our theoretical position and results are generally in accord with those of the aforementioned studies by Mayer and his colleagues. However, the results of Experiment 3 seem to contradict Moreno and Mayer (2002), who found that when no visual diagrams were presented, concurrent presentations of the same auditory and visual text produced better results than did auditory-only text. The inconsistency of results may be resolved by considering the size of textual segments that learners process continuously without a break. In Experiment 3 the text was continuously presented to participants as a single large chunk (of around 350 words) from the beginning to the end, without any breaks. The process of referencing, reconciling, and integrating visual and auditory components of such a large amount of information might have imposed a heavy working memory load and thus inhibited learning. Cognitive load might be expected to be reduced when the text is presented as many consecutive smaller segments with appropriate breaks between them. Smaller segments should allow participants to consolidate partial mental models constructed from each segment of the text before moving to the next one. Such formats of presentation were used in Moreno and Mayer's (2002) experiments.

When text is presented in small, easily managed sequential portions with sufficient temporal breaks between them, a concurrent presentation of identical written and auditory material might not cause deleterious effects on learning, as compared with uninterrupted presentation of the same text as a whole unit. Processing redundant information may overload working memory when learners are dealing with intrinsically complex information, and uninterrupted presentation of long textual descriptions could contribute to this complexity by forcing learners to relate and reconcile many elements of auditory and visual information within a limited time.

The Pace of Instructions

The effects in our experiments were obtained under conditions in which instruction time was limited and preset (based on the length of audio narrations). This conclusion parallels that of a recently reported study by Tabbers et al. (2001), who showed that replacing on-screen text with audio was effective only when the pace of instruction was set by the time of the narration and students had no control over the pacing. When students themselves determined the pace of the instruction, there were no differences between the instructional formats.

Implications for Instructional Design

The current findings have clear applications for instructional design and, in particular, multimedia computer-based instruction. Although the present results are part of a much wider picture, they can be considered from a simple instructional design perspective. The common instructional procedure (particularly in multimedia instruction) of presenting identical spoken and written material simultaneously may need to be avoided, especially in conditions of limited instruction time or system-controlled pacing of instruction.

If identical spoken and written material needs to be presented in such conditions, on the evidence of the current experiments, they should be presented nonconcurrently. Extended sections of text should be partitioned into small, logically completed segments (with time breaks between them). If only text-based information is to be presented, with no related graphical information, then a single-method presentation (either visual or auditory, depending on learner choice) may be preferable to simultaneous presentation of text. If auditory and written materials are not identical (e.g., using some written text as an advance organizer or outline of a spoken presentation), simultaneous presentation might still be beneficial. Further, based on the contrasting conditions that lead to the split-attention and redundancy effects, it can be suggested that information that needs to be integrated in order to be comprehended should be presented simultaneously, whereas information that does not need to be integrated should not be presented simultaneously.

Further work needs to be carried out not only to verify the current findings but also to establish more specific instructional guidelines for the optimal size of textual information. In addition, the reported experiments were conducted with learners who were relative novices in the relevant domains. Effects of increasing levels of expertise need to be investigated in longitudinal studies similar to those reported in Kalyuga et al. (2003).

ACKNOWLEDGMENTS

The research reported in this paper was supported by grants from the Australian Research Council to Slava Kalyuga, Paul Chandler, and John Sweller. We wish to thank our collaborating partners, Email Ltd. and the Broken Hill Proprietary Company Ltd. (BHP), for their ongoing support and, in particular, Richard Winter from Email and Eddy Gosek from BHP. We also express our appreciation to Paul Ginns for his help with recording audio instructional materials.

REFERENCES

Baddeley, A. (1998). *Human memory.* Boston: Allyn & Bacon.

Beccue, B., Vila, J., & Whitley, L. K. (2001). The effects of adding audio instructions to a multimedia computer-based training environment. *Journal of Educational Multimedia and Hypermedia, 10,* 47–67.

Brunken, R., Plass, J., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38,* 53–61.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293–332.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245.

Hegarty, M., Quilici, J., Narayanan, N. H., Holmquist, S., & Moreno, R. (1999). Multimedia instruction: Lessons from evaluation of a theory-based design. *Journal of Educational Multimedia and Hypermedia, 8,* 119–150.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38,* 23–31.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1–17.

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.

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

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

Mayer, R. E., & Anderson, R. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84,* 444–452.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93,* 390–397.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93,* 187–198.

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. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43–52.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86,* 389–401.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43,* 31–43.

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

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94,* 156–163.

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.

Najjar, L. (1996). Multimedia information and learning. *Journal of Educational Multimedia and Hypermedia, 5,* 129–150.

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., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design. *Educational Psychologist, 38,* 1–4.

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

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

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

Paivio, A. (1990). *Mental representations: A dual-coding approach.* New York: Oxford University Press.

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

Schneider, W., & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 53–119). New York: Academic.

Sweller, J. (1999). *Instructional design in technical areas.* Melbourne: Australian Council for Educational Research.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). San Diego: Academic.

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

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

Tabbers, H. K., Martens, R. L., & Van Merrienboer, J. J. G. (2001). *The modality effect in multimedia instructions.* Paper presented at the Annual Meeting of the Cognitive Science Society, Edinburgh.

Tergan, S. (1997). Misleading theoretical assumptions in hypertext/hypermedia research. *Journal of Educational Multimedia and Hypermedia, 6,* 257–283.

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 Someren, M. W., Reimann, P., Boshuizen, H. P. A., & de Jong, T. (Eds.). (1998). *Learning with multiple representations.* Amsterdam: Pergamon.

Slava Kalyuga is a senior research officer at the Educational Testing Centre, University of New South Wales, where he received his Ph.D. in education in 1998.

Paul Chandler is an associate professor of education at the University of New South Wales, where he received his Ph.D. in education in 1993.

John Sweller is a professor of education at the University of New South Wales. He received his Ph.D. in psychology in 1972 from the University of Adelaide, Australia.

Date received: September 16, 2002
Date accepted: January 12, 2004