

# Strategies to reduce the negative effects of spoken explanatory text on integrated tasks

Anne-Marie Singh<sup>1</sup> · Nadine Marcus<sup>2</sup> · Paul Ayres<sup>1</sup>

Received: 6 July 2016 / Accepted: 17 November 2016 / Published online: 23 November 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** Two experiments involving 125 grade-10 students learning about commerce investigated strategies to overcome the transient information effect caused by explanatory spoken text. The transient information effect occurs when learning is reduced as a result of information disappearing before the learner has time to adequately process it, or link it with new information. Spoken text, unless recorded or repeated in some fashion, is fleeting in nature and can be a major cause of transiency. The three strategies investigated, all theoretically expected to enhance learning, were: (a) replacing lengthy spoken text with written text (Experiments 1 and 2), (b) replacing lengthy continuous text with segmented text (Experiment 1), and (c) adding a diagram to lengthy spoken text (Experiment 2). In both experiments on tasks that required information to be integrated across segments, written text was found to be superior to spoken text. In Experiment 1 the expected advantage of segmented text in reducing transitory effects was not found. Compared with written continuous text the segmented spoken text strategy was inferior. Experiment 2 found that adding a diagram to spoken text was an advantage compared to spoken text alone consistent with a multimedia effect. Overall, the results suggest that spoken text is a cause of the transient information effect, which can be best avoided by substituting written text for spoken text on tasks that require integration of information.

**Keywords** Transient information effect · Cognitive load · Spoken explanatory text · Multimedia effects

---

✉ Paul Ayres  
p.ayres@unsw.edu.au  
Anne-Marie Singh  
anne-marie.singh@unsw.edu.au  
Nadine Marcus  
nadinem@unsw.edu.au

<sup>1</sup> School of Education, University of New South Wales, Sydney, Australia

<sup>2</sup> School of Computing Science and Engineering, University of New South Wales, Sydney, Australia

Spoken explanatory text is used in many learning environments. Whether it is a direct exposition, or a voice-over as part of a multimedia presentation, spoken narrative is a very common strategy to provide information and explanations. However, spoken explanatory text can also have a negative effect on learning due to transient information effects (Leahy and Sweller 2011). Spoken text by its nature can disappear before the learner has time to adequately process it, and link it with other or new information. To retain and integrate such information without additional aids, can consume valuable working memory resources leading to a loss of learning- a situation called the transient information effect (Sweller et al. 2011).

Among many strategies proposed to deal with transient information, replacing spoken text with identical written text, or replacing continuous text with segmented text, have both been shown to be effective (Singh et al. 2012). Singh et al. found that for highly sequential independent sections of information both strategies were effective. However, for information that is more connected, and requires integration across segmented information, there are theoretical grounds for assuming that segmentation would be less effective. Hence, the first aim of this study is to show that for integration tasks, replacing spoken text with written text is a more effective strategy than segmenting the spoken information.

The second aim of the study is to investigate the impact of adding diagrams to spoken text for integration tasks. It is a well-researched phenomenon that adding diagrams to text increases learning effectiveness and is called the multimedia effect (Mayer 2014; Mayer and Pilegard 2014). However, again there are theoretical grounds for assuming that adding diagrams to spoken text does not necessarily make it more effective, due to transient information effects. Hence, the study also investigates whether adding a diagram assists learning from spoken text, or is simply replacing it with written text more effective.

## Literature review

### The transient information effect

Much of the research into the transient information effect has been conducted when comparing instructional animations with static presentations. In many studies the expected advantage of animations over statics has not been found (see Kühl et al. 2011; Tversky et al. 2002). Furthermore there are examples where statics have been found to be superior to animations (Castro-Alonso et al. 2014; Mayer et al. 2005). One explanation for this lack of effectiveness is that animations contain transitory images that constantly change. If previous information that vanishes from the display has to be remembered and integrated with current images, then such additional processing may tax the learner's limited working memory (Miller 1956; Cowan 2001). From a cognitive load theory perspective any additional cognitive effort spent on the unnecessary processing of information will direct cognitive resources away from learning (Sweller et al. 2011, 1998). In the case of dynamic displays there may not be sufficient time for the learner to adequately process this information before it disappears from view. In contrast, static pictures (such as diagrams in a book) are more permanent and can place less demand on the temporary storage and processing of information, as they can be restudied a number of times without disappearing if necessary.

The research evidence suggests that transitory information may have a negative impact on the effectiveness of dynamic representations; however, the most common form of

transitory information is spoken text. Thus the same issue that is found with instructional animations can be found with spoken explanatory text. By its very nature, spoken text is transitory. Unless it is recorded in some fashion once spoken it disappears. Hence, if multiple segments or sub-segments of information have to be integrated together to learn about a specific concept, then a considerable amount of working memory may be required to deal with these demands, instead of focusing on learning. In contrast, the permanency of written text allows readers to revisit the same information a number of times if needed, requiring less working memory resources and enabling more to be directed to learning.

Evidence that spoken narration impacts on learning has been demonstrated by Leahy and Sweller (2011), who investigated the modality effect with elementary school students learning how to interpret temperature–time graphs. The modality effect is a well-known effect in multimedia research and occurs when spoken text plus diagrams facilitate more learning than identical written text plus diagrams (Low and Sweller 2014). Leahy and Sweller (2011) found that for fairly short explanatory text a modality effect occurred; however, for more lengthy explanatory text a reverse modality effect occurred (written text plus diagrams was superior to spoken text plus diagrams). Similarly, Wong et al. (2012) found the same effects using temperature–time graphs. In both studies the usual advantage of combining spoken text and visual information stored in both working memory channels (see Baddeley and Hitch 1974; Low and Sweller 2014) was negated with lengthy text because of the difficulty dealing with the fleeting spoken information.

### Studies comparing written text with spoken text

In addition to investigating the impact of spoken text on the modality effect, a number of studies have directly compared the effectiveness of written text with spoken text. For example, Singh et al. (2012) found written text to be superior to spoken text in learning about highly sequential self-contained information (the 8 steps to passing a government bill). Learning advantages for written text have also been demonstrated in the field of communications (e.g., Byrne and Curtis 2000; Furnham and Gunter 1985, 1989). Byrne and Curtis (2000) observed that the written medium was a more efficient way to present complex information related to health. Using news narratives, Furnham and Gunter (1985, 1989) found that learners could recall more from a sequence of violent and non-violent news stories material when information was presented in a written only format compared to spoken-plus-diagrams or a spoken-only medium.

A number of explanations have been provided to explain the advantages of using written text. Firstly, readers can choose their desired reading rate by slowing down or increasing their reading rate to facilitate more understanding (Byrne and Curtis 2000). Secondly, readers have the ability to re-read sections of the text, which allows the reader more time to study difficult and/or confusing text passages (Danks and End 1987; Furnham et al. 2002). Thirdly, the reader has the ability to skip or skim text passages which are difficult or not necessary to understanding the topic, allowing the learner to focus on relevant or complex elements of the task (Bazerman 1985). Such explanations highlight that when reading the processing strategies are applied by the learner, and not administered and controlled by the instructor or instructional materials as with spoken text. Such arguments provide feasible explanations for the advantages of written text over spoken text without directly pointing out the negative aspects of spoken text caused by its transient nature.

## Overcoming transient information by segmentation

As well as exchanging spoken text with written text other strategies have been used to overcome the negative effects of transient information, particularly in animated and other multimedia learning environments. These include learner control (e.g. Crooks et al. 2012; Hasler et al. 2007; Mayer and Chandler 2001; Mayer et al. 2003) and signalling strategies (e.g. de Koning et al. 2009; Mautone and Mayer 2001). However, a common strategy, and most relevant to the current study, is segmentation. Segmentation occurs when information is presented in smaller sections with an identifiable beginning and end point (Spanjers et al. 2010), which manages transience by reducing the amount of information that needs to be processed simultaneously at one given time.

An increasing number of studies have found that higher learning outcomes arise from animations that have been segmented into smaller sections. In comparison to non-segmented (continuous) animations, supporting evidence in favor of segmented animations has been found in a number of domains, such as learning about the formation of lightning (Cheon et al. 2014; Hassanabadi et al. 2011; Mayer and Chandler 2001), paper folding tasks (Wong et al. 2012), remembering soccer sequences (Khacharem et al. 2013), knot tying (Wong et al. 2009), and calculating probability (Spanjers et al. 2012, 2011).

In a direct comparison of spoken and written text without other multimedia aids, Singh et al. (2012) found that not only was written text superior to spoken text (reported above), but also segmented text was superior to continuous text. In this study there was also a significant interaction indicating that segmentation was an advantage for spoken text but not for written text.

To explain how segmentation supports learning, Spanjers et al. (2010) proposed two possible reasons that both stem from a key feature of segmentation, in that it is created by inserting pauses. Firstly, pauses provide learners with additional time to process and integrate the information received in previous segments without the need to simultaneously attend to newly incoming information. Therefore pauses provide learners not only extra time on task but also help learners deal with transitory information (Spanjers et al. 2012). Secondly, segments present information in meaningful chunks, which provide cues for learners enabling them to find boundaries within the information presented without additional searching processes (Schnotz and Lowe 2008; Wouters et al. 2008). Whereas the first reason supports an argument for dealing with transitory information, the second suggests that cuing or signaling strategies (see van Gog 2014) underpin the effectiveness of segmentation. Both reasons are consistent with reducing unnecessary cognitive load and therefore enhancing learning (Sweller et al. 2011).

Segmenting information can also cause other effects. Research from the field of memory has shown that segmenting information creates event boundaries, and crossing these boundaries can have both positive and negative impact on recall. Radvansky (2012) reports that encountering event boundaries can lead to more processing and effort required to maintain information, and hence increases cognitive load. Specifically, it can lead to memory loss from previous segments as boundaries are crossed. Radvansky (2012) also reports that if the segments are connected causally, then information is more likely to be remembered. Regarding the Singh et al. (2012) study that featured highly sequential and causal information (the steps in a government bill), segmentation could have been a helpful memory aid as well as helping deal with transitory information. Without such causal connections segmentation may be less effective.

## Multimedia effects

Whereas learning can occur from spoken or written text alone there are many advantages from using multimedia materials. Simply adding a diagram or picture to text can have a positive advantage. The multimedia effect occurs when learning from text plus diagrams is found to be superior to learning from text alone (Mayer 2014; Mayer and Pilegard 2014). It is usually argued, based on dual-coding theory (see Paivio 1986), that using both text and diagrams forms a stronger mental representation of the learning concept than text alone (Fletcher and Tobias 2005; Mayer and Fiorella 2014).

Another well-known effect, called the modality effect, occurs when spoken text and diagrams leads to greater learning than written text plus diagrams (Low and Sweller 2014; Mousavi et al. 1995). Like the multimedia effect it is has been replicated many times and has its theoretical underpinnings based on dual-processing systems in working memory (see Baddeley and Hitch 1974). Under many conditions, in conjunction with diagrams or pictures, spoken text has a clear advantage over written text because less working memory resources are required to listen to spoken text while examining a diagram than reading text and trying to match up information in a diagram (a form of split attention, see Ayres and Sweller 2014).

However, recent research has shown that lengthy spoken text can cause a reverse modality effect, where spoken text plus diagrams leads to inferior learning compared to written text plus diagrams (Crooks et al. 2012; Leahy and Sweller 2011; Schüler et al. 2012; Wong et al. 2012). Consistent with a transient information argument, lengthy spoken text is more difficult to process and therefore raises cognitive load sufficiently to overcome the normally positive effects of spoken text plus diagrams.

The multimedia and modality effects generally support the view that adding a diagram to spoken text would increase learning compared to spoken text alone. But as discussed above, the inclusion of spoken text does not necessarily guarantee the best learning outcomes because of its transient nature. In contrast, written text may have more advantages because of its more permanent nature.

## The current study

The study investigates the effectiveness of three strategies that could overcome the negative effects of transient information created by spoken explanatory text. The first strategy replaces spoken text with written text; the second sub-divides spoken text into more manageable segments; and the third examines how adding a diagram to spoken text improves its effectiveness. The main focus of this study is to examine how such strategies deal with information that needs to be extracted and integrated across segments (integration tasks). However, in order to serve as a comparison and offer potential further insights, tasks are also designed that require information to be extracted from individual segments only (segment tasks). Hence, the study will investigate whether there are any significant differences in effectiveness between the three nominated strategies on both segment and integration tasks.

## Study hypotheses

### *Hypotheses developed for Experiment 1*

Previous research has found that written continuous text is superior to spoken continuous text for highly segmented information (see Singh et al. 2012). We argue that this effect can be extended to integration tasks as well. Both segment and integration tasks will be more difficult in a spoken text medium, due the highly transitory nature of spoken text, which causes high cognitive load. Hence it was predicted that:

For integration tasks, written continuous text will lead to more learning with less cognitive load than spoken continuous text (Hypothesis 1a).

For segment tasks, written continuous text will lead to more learning with less cognitive load than spoken continuous text (Hypothesis 1b).

As reported above, replacing spoken explanatory text with written text, or segmenting spoken text, are both effective learning strategies. Therefore, to compare these two strategies the pros and cons of each strategy are first considered. As reported above, the main strength of written text is that it is less transitory than spoken text, as it is more permanent and can be revisited if necessary a number of times in the given time frame. Hence cognitive load is reduced compared to spoken text. A potential negative compared with segmentation is that information boundaries are less noticeable, which may require more working memory resources to locate them, although some empirical evidence does not support this impact (Singh et al. 2012). Furthermore, for tasks that require information to be integrated across segments, locating boundaries may be less problematical.

Segmentation has the possible advantages of dealing with transitory information through pauses and more clearly defined boundaries (Spanjers et al. 2010). However, there are some potential negatives. If as Spanjers et al. (2012) suggest, a significant characteristic of segmentation is that information boundaries become more salient, then a potential danger is that segmenting may cause a greater focus to be made on boundaries and information contained within them. For tasks that require information to be integrated across boundaries, a focus on segments may have a negative impact. Even though pausing may reduce the impact of transitory information, spoken text is still transitory, even when segmented, which may reduce its overall effectiveness compared with written text which is always more permanent. Hence, based on the above argument it is predicted that:

For integration tasks, written continuous text will lead to more learning with less cognitive load than segmented spoken text (Hypothesis 2a)

In the case of segment tasks, a different hypothesis can be generated. As Singh et al. (2012) found both segmenting spoken text and substituting written text for spoken text led to superior learning compared to continuous spoken text. Therefore, it is predicted that:

For segment tasks, no differences in learning and cognitive load will occur between written continuous text and segmented spoken text (Hypothesis 2b)

In comparing spoken continuous text and spoken segmented text, Singh et al. (2012) found that for segment tasks, segmenting spoken text was an advantage. However, for integration tasks, we argue that segmentation is less likely to be effective because of the potential impact of integrating across boundaries, as described earlier. Nevertheless segmenting

highly transient spoken text may still provide some advantage in reducing cognitive load overall. Therefore we predict that there would be no difference between the two strategies for integration tasks. The following hypotheses are made:

For integration tasks, no differences in learning and cognitive load will occur between segmented spoken text and continuous spoken text (Hypothesis 3a).

For segment tasks, segmented spoken text will lead to more learning with less cognitive load than continuous spoken text (Hypothesis 3b).

### *Hypotheses developed for Experiment 2*

In Experiment 2 a diagram is added to the text in a 2 (no diagram vs. diagram)  $\times$  2 (spoken text vs. written text) design, in order to investigate potential multimedia effects. As described in the literature review the multimedia effect is well established and occurs when learning from text plus diagrams is found to be superior to learning from text alone (Mayer 2014; Mayer and Pilegard 2014). It is expected that this effect will occur for both segment and integration tasks. Therefore it is predicted that:

For integration tasks, text (both spoken and written) plus a diagram will lead to more learning with less cognitive load than text (both spoken and written) alone (Hypothesis 4a).

For segment tasks, text (both spoken and written) plus a diagram will lead to more learning with less cognitive load than text (both spoken and written) alone (Hypothesis 4b).

Similarly, a second main effect can be predicted for both types of tasks based on the evidence that written text is superior to spoken text.

For integration tasks, written text will lead to more learning with less cognitive load than spoken text (Hypothesis 5a).

For segment tasks, written text will lead to more learning with less cognitive load than spoken text (Hypothesis 5b).

### **Design features of the study**

To test these hypotheses a number of important design features were considered, including the length of a segment, how and where to insert a pause, and how long a pause should last.

#### *Length of the segment*

Schüler et al. (2011) found that the advantages of spoken text plus diagrams disappeared when five or more sentences were used per visualization. Therefore segments in this study were limited to two to four sentences of information.

#### *Segment boundaries*

An event boundary occurs when there is a physical or structural change in the text. The location of where segments start and finish determine where to insert pauses, ultimately creating segments within the text (see Spanjers et al. 2010). Therefore to ensure the text

was presented in an optimum segmented format, segments were designed to contain information on a complete concept using whole sentences.

### *Duration of a pause*

Segmentation increases total study time by inserting pauses throughout the text. The additional time required to process information should be long enough so that learners can process and link information to prior knowledge without having to attend to new incoming information, nor being excessively long as learners may forget the information. Pauses ranging from 2 s (e.g., Khacharem et al. 2013; Spanjers et al. 2012, 2011) to 30 s (Wong et al. 2012) have proven to be successful, however Spanjers et al. (2012) suggested that 2 s is probably too short for processes such as reflection during a pause. Therefore a 5 s pause was chosen in this study to ensure that there was adequate time to process information between segments. This pause length was chosen as it fitted the guidelines suggested above, and had also been used successfully by Singh et al. (2012) in a study involving spoken segmented text.

### *Types of tasks*

An important focus of the study was to test the hypotheses on information that had to be integrated across segments. The *integration-retention* tasks were specifically designed for this purpose, as they required participants to integrate information from a number of different segments. Similarly, the transfer task was an extended response style question that required participants to use the total information presented to make an evaluation of a particular problem. In addition, in order to serve as a comparison, tasks were also designed that required information to be extracted only from individual segments (segment tasks).

## **Experiment 1**

In this experiment three groups were compared: Written continuous text, spoken continuous text, and spoken segmented text. The learning tasks were based on high school economics. Hypotheses 1a, 1b, 2a, 2b, 3a, and 3c were tested in this experiment.

## **Method**

### *Participants*

Forty-six boys in grade 10 (15- and-16 years old) studying commerce from one Sydney high school participated in the study. Participants were invited to participate in the study based on their enrolment in the 200-h New South Wales State Commerce course (NSW Board of Studies 2003) designed for students this age. Prior to the study commencing, a cohort of students was randomly allocated to one of the three experimental groups. Due to variations in attendance on the day of testing group numbers differed slightly: written-continuous ( $N = 17$ ), spoken-continuous ( $N = 15$ ) and spoken-segmented ( $N = 14$ ). Participants were deemed to have an appropriate level of knowledge given that they had completed half of the commerce course in the previous year and had learnt fundamental



economic concepts. Nonetheless, it was still expected that participants would find the tasks challenging.

### *Materials and procedure*

The experiment consisted of three phases in the following sequence: an instructional learning phase (6 min and 40 s, including a 60 s pause between presentations), a retention test phase (20 min), and a transfer test phase (20 min). After both the retention and transfer tests, participants were required to rate their mental effort to collect an index of cognitive load. All materials were distributed and collected after each phase. No guidance or feedback was given to the participants throughout the experiment, which was completed during a scheduled Commerce lesson of 110 min. Students completed the experiment in groups according to their randomized interventions under strict supervision.

### *Initial procedure*

Prior to commencing the study all participants were provided with an auditory explanation outlining a general overview of the study. Participants were informed of the topic (“the economic cycle”) and given a brief description of each phase.

### *Instructional learning phase*

The materials studied during the learning phase included an explanatory text describing key terms in addition to explaining the natural fluctuations in the economy (periods of expansion and contraction). The text explained how: total output, consumer spending, inflation, wage rates, interest rates and unemployment determine the current stage of the economic cycle. Furthermore, the text explained the causes and implications of these changes on consumer, business and government spending. Information was carefully divided into 8 segments (S1–S8) resulting in slight variations of word numbers (w) in each segment (S1 = 36w, S2 = 35w, S3 = 31w, S4 = 41w, S5 = 34w, S6 = 44w, S7 = 36w and S8 = 41w). Each segment consisted of 2–4 full sentences. Each segment contained information about a discrete concept.

To compare the three instructional groups, all groups were presented with identical information for the same duration. In order to achieve this equivalence, the spoken-continuous format was used as the basis for constructing the other instructional groups. It took the speaker 130 s to read through the explanatory information at a steady continuous pace. The text was divided into 8 segments. Pauses were used as a device for segmenting the spoken text. A 5 s pause was inserted after each segment to provide clear boundaries. Therefore, 170 s ( $130 + 8 \times 5$ ) became the time frame for each presentation including 40 s of pauses.

The spoken-continuous group listened to a female voice recording of the spoken text (130 s), followed by a pause of 40 s at the end of the text to control learning time. The spoken-segmented group listened to the same 130 s voice recording, however the text was divided into 8 segments followed by a 5 s pause after each segment. All auditory information was heard on individual laptops using headphones. Information was uploaded, controlled and monitored via the school’s intranet system (iLearn). Participants were allowed to start the recording upon instructions from the supervisor. In addition participants were instructed prior to the experiment and reminded before the learning phase that

they could not stop, rewind or replay the recording once it had started. Once participants started the recording they heard the presentation for the first time, followed by a 60 s pause, and then they automatically heard the presentation again.

The written-continuous group received a typed word document containing the same information provided to the spoken groups. The written information was presented on one A4 page using size 12 Times New Roman Font with single line spacing. The text was presented as a single continuous paragraph. The written group listened to a female voice recording, which instructed the participants to start and stop reading. Participants were given 170 s to read through the text. Following the full 170 s presentation, the 60 s pause started and the participants were instructed to turn the page over to prevent additional learning. After the pause, the recording instructed participants to turn the page over and start reading again.

### *Retention test phase*

To test the effectiveness of the instructional condition participants were required to answer 12 questions. Questions 1–8 required a direct recall of information from individual segments. For example, the answer to Question 1 could be found in Segment 1, Question 2 could be found in Segment 2 and so forth. Two lines were provided after each question for participants to write their answers. Questions 9–12 required participants to integrate information from multiple segments across the text. For example, Question 10 stated ‘Explain the impact on consumer confidence during an expansionary phase’. Participants needed to recall information from S1, S2, S3, S4 and S8. Four empty lines were provided per question. All questions and answer lines were printed over two A4 pages.

### *Transfer test phase*

To test if knowledge was processed at a deeper level, participants were presented with one complex extended response style question. To provide some guidance in answering the transfer question, it included a statement followed by a question. The transfer question read as (a) ‘Humans are largely affected by the optimism or pessimism of others: most of us tend to be upbeat together and willing to spend; or we share the pervasive gloom and are thrifty’; (b) ‘Evaluate the impact the above statement has on business spending and government spending if the Australian economy is in a bust’. Participants were required to use the information learnt during the instructional phase to make an evaluation on the statement and illustrate their response with supporting arguments. Forty-five lines were provided and printed on two A4 pages. Instructions were printed on the front page of the transfer booklet.

### *Cognitive load measure*

To get a measure of cognitive load a 1-item self-rating scale (mental effort) was used (see Paas 1992). This instrument provides an accurate and non-evasive indicator of cognitive load, which has been used successfully in many studies (see van Gog and Paas 2008; Sweller et al. 2011). At the end of the retention and transfer phases, participants were instructed to circle their level of mental effort invested (‘1 = extremely low, 2 = very low, 3 = slightly low, 4 = neither low or high, 5 = slightly high, 6 = very high and 7 = extremely high’). This scale was printed on the last page of each test booklet.

### Scoring of tasks

The segment-retention task (questions 1–8) phase was marked out of thirteen and half marks. Marks were allocated according to key terms consistent with the NSW Board of Studies marking guidelines (NSW Board of Studies 2003) for this task: define (i.e. state the meaning and identify essential qualities, 2 marks), outline (i.e. sketch in general terms/indicate the main features of..., 2 marks) and identify (i.e. recognize and name, 1 or 1.5 marks). A maximum of 2 marks were awarded for Questions 1, 2, 3, 6 and 8. For partially correct answers, participants were awarded 1 mark. All other answers were awarded 0 marks. A maximum of 1.5 marks was awarded for Question 4. Participants were required to list 3 points. For each correct point referenced, a half mark was awarded. A maximum of 1 mark was awarded for Questions 5 and 7, where participants were required to list two points. For each correct point referenced, a half mark was awarded.

The integration-retention test (Questions 9–12) was awarded a maximum of 4 marks per question. For each question a list of possible 6 points were identified according to the State curriculum guidelines (NSW Board of Studies 2003). To achieve the maximum 4 marks participants needed to include a minimum of 4 characteristics and relate cause and effect between each point. Participants were awarded 3 marks if they described 3 characteristics, 2 marks for 2 characteristics, and 1 mark for 1 characteristic. If participants attempted the question and partially made reference to one of the points from the list they were awarded a half mark. All other responses were awarded 0 marks.

The transfer test phase was allocated a maximum of 20 marks based on list of 10 key points generated by two experienced teachers. All 10 points in the list needed to be referenced to achieve maximum marks.

Cronbach alphas were calculated for the segment-retention task ( $\alpha = 0.658$ ) and the integration-retention task ( $\alpha = 0.802$ ), providing acceptable internal consistency, especially for the latter task. A Cronbach alpha was not calculated for the transfer task as it was allocated a single mark. All scores were converted to percentages to allow easy comparisons between the 3 tasks.

## Results

Group mean test scores are shown in Table 1 and mean mental effort scores in Table 2. For each measure a 1-way ANOVA was conducted with follow up Bonferroni tests (adjusted) when significant overall results were found. For the *segment-retention task* (Questions 1–8), there was no significant difference between the 3 groups ( $F < 1$ , ns.). For the *integration-retention task* (Questions 9–12), there was a significant difference between the groups  $F(2, 43) = 7.35$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.255$ . Follow up tests indicated that the written continuous text group had significantly higher scores than the spoken continuous text group ( $p = 0.027$ ) and the spoken segmented text group ( $p = 0.002$ ). No other comparisons were

**Table 1** Group means (and SDs) of test scores (%) for Experiment 1

Groups	Segment-retention	Integration-retention	Transfer
Written-continuous	71.5 (21.8)	56.3 (26.3)	67.4 (25.5)
Spoken-continuous	65.9 (21.4)	35.8 (16.8)	63.2 (18.4)
Spoken- segmented	62.7 (18.9)	28.6 (17.6)	53.2 (18.0)
Overall total	67.0 (20.7)	41.2 (23.8)	61.7 (23.8)

**Table 2** Group means (and SDs) for mental effort scores in Experiment 1

Groups	Test mental effort	Transfer mental effort
Written-continuous	3.65 (1.22)	3.94 (1.75)
Spoken-continuous	4.80 (1.08)	4.33 (1.29)
Spoken- segmented	4.25 (0.94)	4.72 (1.49)
Overall total	4.21 (1.09)	4.30 (1.52)

significant. For the *transfer task*, there was no significant difference between groups,  $F(2, 43) = 1.76$ ,  $p = 0.185$ ,  $\eta_p^2 = 0.076$ . For the *test mental effort* measure there was a significant difference between groups,  $F(2, 43) = 4.43$ ,  $p = 0.018$ ,  $\eta_p^2 = 0.171$ . Follow up tests indicated that the written continuous text generated significantly lower ratings than the spoken continuous text ( $p = 0.015$ ). No other comparisons were significant. For the *transfer mental effort* measure, there was no significant difference between groups,  $F(2, 43) = 2.30$ ,  $p = 0.113$ ,  $\eta_p^2 = 0.096$ .

## Results summary and discussion

The main aim of this experiment was to examine how effective it was to replace spoken continuous text with written continuous text or spoken segmented text for tasks that required information to be extracted and integrated across segments. Support was found for Hypothesis 1a in that for the *integration-retention task* (Questions 9–12) significantly more learning occurred with the written continuous text compared to the spoken continuous text, with reduced mental effort (cognitive load). Some support was also found for Hypothesis 2a in that for the *integration-retention task* significantly more learning occurred with the written continuous text compared to the spoken segmented text. No evidence was found that spoken segmented text led to superior learning with reduced cognitive load compared with spoken continuous text (Hypothesis 3a).

There were no significant differences on the *segment-retention task* (Questions 1–8) for any of the three comparisons. Hence no support was found for Hypothesis 1b (written continuous text superior to spoken continuous text) or Hypothesis 3b (segmented spoken text superior to continuous spoken text) on this segment task. However, Hypothesis 2b (no differences between written continuous text and spoken segmented text) was supported.

Furthermore, no significant effects were found for the transfer task, which did require the integration of information. However, the total mean scores on the segment-retention tasks and transfer tasks were 67.0 and 62% respectively, which were considerably higher than the integration-retention task (41%), suggesting that significant effects may have been more likely on the difficult task. It is notable that the Spoken-segmented group did not score higher than the Spoken-continuous group on any of the three tasks. Hence, no evidence emerged that segmenting spoken text was an advantage in the domain, which contained interrelated information that required integration.

## Experiment 2

The results for the integration-retention task in Experiment 1 suggests that presenting complex information in a spoken-only format is a less effective medium for learning, which is consistent with the assumption that spoken text is a source of transient information (Sweller et al. 2011), which can be reduced by replacing it with more permanent

written text. Segmenting spoken text was not found to have an advantage, and was significantly inferior to written text.

The aim of Experiment 2 was to broaden the investigation into strategies that reduce the transitory effects of spoken text by including a multimedia approach (adding an instructional diagram). Four experimental groups were compared in a 2 (no diagram vs. diagram)  $\times$  2 (spoken text vs. written text) design. Accordingly Hypotheses 4a, 4b, 5a, and 5b, were tested using 2  $\times$  2 ANOVAs.

## Method

### *Participants*

Seventy-nine grade 10 boys (15–16 years) from one Sydney high school participated in the study. Participants were invited to participate in the study based on their enrolment in the 200-h State Commerce course (NSW Board of Studies 2003). Prior to the study commencing each expected participant was randomly allocated to one of four experimental groups. Due to absences on the day of testing, group numbers varied slightly as follows: the written-only group ( $N = 21$ ), the written-plus-diagram group ( $N = 18$ ), the spoken-only group ( $N = 19$ ), and the spoken-plus-diagram group ( $N = 21$ ).

### Materials and procedure

The design of the study closely followed the materials and procedure of the previous experiment and included the same content on the economic cycle. Again there were three phases: A learning phase (6 min: including a 60 s pause between presenting the materials), retention test (20 min) and transfer test (20 min). All participants were given identical tests. The only difference between the four groups occurred during the learning phase, where participants studied the content according to the instructional condition that they were assigned to. All materials were distributed and collected after each phase. No guidance or feedback was given to the participants throughout the experiment, which was completed during a scheduled commerce lesson (110 min).

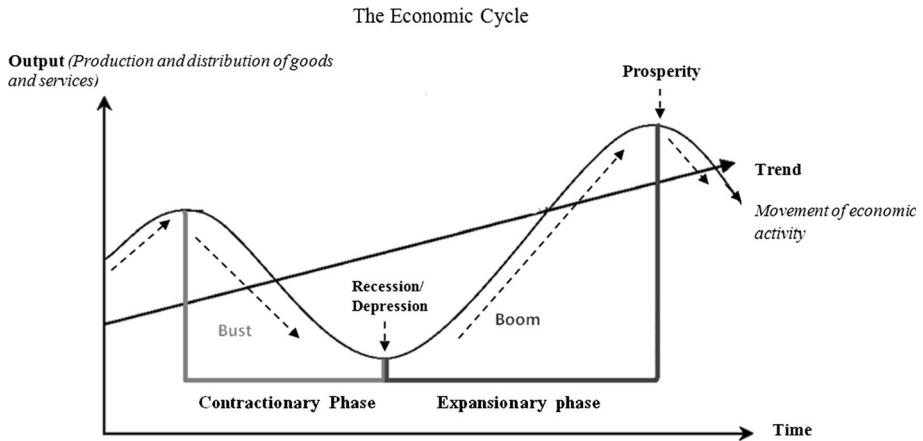
### *Pre-instructions*

Information given prior to the commencement of the study was identical to Experiment 1.

### *Instructional conditions*

The same materials and procedure in Experiment 1 were again used with the following differences. Previously the spoken-segmented group was used as a benchmark for all other conditions. It took 130 s to read through the explanatory information at a steady continuous pace, the text was then divided into 8 segments, and a 5 s pause was inserted after each segment. Therefore, 170 s became the time frame for each presentation. As the effects of segmentation were not tested in the current experiment the 40 s pauses were removed, hence 130 s became the time frame for each presentation.

The spoken groups listened to a 130 s female voice recording of the spoken text. The spoken-plus-diagram group listened to the same 130 s voice recording; however they were also given a diagram of the economic cycle, which was printed at the top of a single A4



**Fig. 1** Diagram of the economic cycle used in experiment 2

sheet of paper (see Fig. 1). The written groups read the written text for 130 s. The written-plus-diagram group read the text and studied the diagram (see Fig. 1) for 130 s. The cycle was then repeated once. Therefore all four groups were given the same time period to study the materials.

The role of the diagram was to give a pictorial overview of the economic cycle and help participants consolidate the explanatory text in some instances. The explanatory text alone was sufficient to answer all questions. However, studying the diagram in isolation would have only provided enough information to answer 5 questions (Q1, Q2, Q3, Q5, and Q7) in the segment-retention task, but not for the remaining 3 questions (Q4, Q6, and Q8). Therefore the diagram alone would not have been sufficient for participants to complete any of the questions in the integration-retention task (Questions 9–12) or the transfer task. Hence explanatory text was required to complete all tasks.

All auditory information was heard on individual school laptops using headphones. Information was uploaded, controlled and monitored via the school's intranet system (iLearn). Participants were allowed to start the recording upon the given instruction.

The written-only group and written-plus-diagram group received an identical typed word document containing the same information that was provided to the spoken groups, however the written-plus-diagram group were provided with the same static diagram of the economic cycle as presented to the spoken-plus-diagram group, positioned underneath the written text. All print information was presented on one A4 page. The groups who received written texts listened to a female voice recording of instructions to start and stop reading. Participants were allowed to start the recording upon a given instruction.

Each group was given a 60 s pause between presentations. Participants in the written-only, the written-plus-diagram and the spoken-plus-diagram groups were instructed to turn the page over during this pause, to prevent additional learning time. After 60 s they were then instructed to turn the page over and start reading. Participants in the spoken-only text group automatically heard the text start after the 60 s pause.

### *Tests and cognitive load measures*

The performance tests used were identical to the previous experiment consisting of a segment-retention task, an integration-retention task, and a transfer task. The same measure

of cognitive load was used (mental effort). All procedures were identical to the previous experiment.

### Scoring of tests

The same scoring system was used again. Cronbach alpha scores were 0.817 on the segment-retention task and 0.844 on the integration-retention task, providing high internal consistency.

## Results

Scores were added and grouped according to the different types of questions asked. Mean performance scores are shown in Table 3. Mean scores per group for mental effort are shown in Table 4.

### Hypotheses 4a and 4b: text versus text plus diagram

For the *Segment-retention task* (Questions 1–8) there was no significant main effect for diagrams,  $F(1, 75) = 1.89, p = 0.202, \eta_p^2 = 0.025$ . For the *Integration-retention task* there was a significant main effect for diagrams,  $F(1, 75) = 5.50, p = 0.022, \eta_p^2 = 0.068$ , where participants studying with diagrams ( $M = 6.60$ ) outperformed participants who did not receive a diagram ( $M = 4.79$ ). For the *Transfer task*, there was a significant main effect for diagrams,  $F(1, 75) = 5.30, p = 0.024, \eta_p^2 = 0.066$ , where participants studying with diagrams ( $M = 13.83$ ) outperformed participants studying without diagrams ( $M = 12.0$ ). For the *Test mental effort* measure, participants studying with diagrams ( $M = 4.68$ ) rated

**Table 3** Group means (and SDs) of test scores (%) for Experiment 2

Groups	Segment-retention	Integration-retention	Transfer
Written	67.2 (30.1)	40.2 (23.7)	66.0 (20.0)
Spoken	68.8 (22.6)	18.6 (19.8)	53.0 (16.8)
Combined	68.0 (26.5)	29.9 (24.3)	59.8 (19.4)
Written + diagram	82.7 (23.6)	50.5 (25.6)	73.8 (22.1)
Spoken + diagram	69.3 (25.9)	33.3 (25.3)	65.2 (18.1)
Combined	75.5 (25.5)	41.3 (26.5)	69.2 (20.2)
Overall total	71.7 (26.1)	35.5 (25.9)	64.4 (20.2)

**Table 4** Group means (and SDs) for mental effort scores in Experiment 2

Condition	Test mental effort	Transfer mental effort
Written	4.71 (1.42)	5.43 (1.17)
Spoken	4.63 (1.67)	5.00 (1.63)
Combined	4.68 (1.53)	5.23 (1.40)
Written + diagram	5.00 (1.19)	4.95 (1.17)
Spoken + diagram	5.33 (0.80)	5.00 (1.41)
Combined	5.18 (1.00)	4.97 (1.18)
Overall total	4.92 (1.31)	5.10 (1.30)

mental effort lower than participants studying without diagrams ( $M = 5.18$ ), which may be considered significant under a 1-tailed test as it was consistent with the direction of the predicted difference,  $F(1, 75) = 2.82$ ,  $p = 0.094$ ,  $\eta_p^2 = 0.04$ . For the *Transfer mental effort* measure there was no significant effect ( $F < 1$ , *ns.*).

#### *Hypotheses 5a and 5b: written text versus spoken text*

For the *Segment-retention task*, there was no significant effect,  $F(1, 75) = 1.62$ ,  $p = 0.3316$ ,  $\eta_p^2 = 0.013$ . For the *Integration-retention task*, there was a significant effect,  $F(1, 75) = 13.1$ ,  $p = 0.001$ ,  $\eta_p^2 = 0.149$ , where participants in the written group ( $M = 7.19$ ) outperformed participants in the spoken group ( $M = 4.21$ ). For the *Transfer task*, there was a significant main effect,  $F(1, 75) = 6.09$ ,  $p = 0.016$ ,  $\eta_p^2 = 0.08$ , where the written groups ( $M = 13.8$ ) performed significantly better than the spoken groups ( $M = 12.0$ ). For both mental effort measures there was no significant effects (both  $F < 1$ , *ns.*).

#### *Interactions between text (written vs. spoken) and diagrams (diagrams vs. no diagrams)*

There were no significant interactions for any of the measures: *Segment-retention task*,  $F(1, 75) = 1.65$ ,  $p = 0.203$ ,  $\eta_p^2 = 0.02$ ; *Integration-retention task* ( $F < 1$ , *ns.*); *Retention test mental effort scores* ( $F < 1$ , *ns.*); *Transfer task*, ( $F < 1$ , *ns.*); and the *Transfer mental effort scores* ( $F < 1$ , *ns.*)

### **Summary of results and discussion**

A number of hypotheses were tested in this experiment. Hypotheses 4a and 4b predicted that adding a diagram to text would be superior to text alone (a multimedia effect; Mayer 2014). Significant differences were found on the integration-retention and transfer tasks supporting Hypothesis 4a. But no significant difference was found on the segment-retention task, and therefore there was no support for Hypothesis 4b. No interactions were found indicating that adding a diagram improved learning outcomes for both written and spoken text. Of particular interest to this study concerned spoken text, and as this result confirms-spoken text can be enhanced by adding a diagram.

Hypotheses 5a and 5b predicted that written text would be superior to spoken text. Again significant differences were found on the integration-retention and transfer tasks supporting Hypothesis 5a. But no significant difference was found on the segment-retention task, and therefore there was no support for Hypothesis 5b. The lack of interactions indicated that for both diagram and no-diagram conditions, learning was superior with written text compared to spoken text. Hence, a reverse modality effect (see Leahy and Sweller 2011) can be identified as written text + diagram ( $M = 50.5$ ) was superior to spoken text + diagram ( $M = 33.3$ ).

Support was found for each hypothesis that predicted condition differences, but only on tasks that required the integration of information. In particular, support was found for each hypothesis on the integration-retention task, for two hypotheses on the transfer task; but no support was found on the segment-retention tasks. No significant results were found for either mental effort measure. It can be concluded that written text has advantages over



spoken text, and adding a diagram to spoken and written text is an advantage compared to spoken and written text alone.

## General discussion

The primary aims of the study were to show that for integration tasks, replacing spoken text with written text was a more effective strategy than segmenting the spoken information, and that adding a diagram to spoken text improved its learning effectiveness. The prediction that more learning would occur from written text than segmented spoken text was tested in Experiment 1. Results on the integration-retention task confirmed this prediction as participants in the written text condition had significantly higher scores with less cognitive load. In addition, a direct comparison between continuous written text and continuous spoken text was also made in Experiment 1, showing a significant advantage for written text. In contrast, comparing segmented spoken text with spoken text revealed no significant differences. Furthermore, written text was also directly compared with spoken text in Experiment 2, finding that written text had a significant advantage on both the integration-retention and transfer tasks. Hence it was concluded that written continuous text was a superior alternative to spoken text, rather than segmenting it.

The second main research aim was to investigate whether adding a diagram to spoken text improved its learning effectiveness for integration tasks. Results from Experiment 2 on the integration task and transfer task indicated an overall multimedia effect (Mayer 2014), which also indicated that adding a diagram to spoken text did lead to superior learning compared to spoken text alone. In addition a reverse modality effect (see Leahy and Sweller 2011) was also found, where for the Integration-retention task, written text plus a diagram was superior to spoken text plus a diagram.

## Theoretical implications

The results indicate that presenting complex information in spoken-only format was detrimental to learning for tasks that required the integration of information across segments. This finding is consistent with the assumption that spoken explanatory text is a source of transitory information, creating high cognitive load which can be reduced by presenting information in a more permanent source (Sweller et al. 2011). For spoken text more working memory resources are used trying to remember previous information and link it with new information, hence leaving fewer resources available for learning. In contrast, written text can be revisited a number of times if necessary in order to link together the various concepts within the text that need to be integrated.

No evidence was found that segmenting spoken text improved learning outcomes for spoken text. In contrast, written continuous text was superior to both spoken text and segmented spoken text. Part of the motivation for making this comparison was to gain insights into the factors underlying the positive effects of segmentation. It is argued by researchers that segmentation can reduce or eliminate extraneous cognitive load associated with transience (e.g. Ayres and Paas 2007; Sweller et al. 2011; Wong et al. 2012) by controlling the way learners' process information (Spanjers et al. 2012). Unlike the present study, a number of multimedia studies in particular have observed that a segmentation strategy improves retention and transfer of knowledge (e.g., Leahy and Sweller 2011; Mayer and Chandler 2001; Mayer et al. 2003; Spanjers et al. 2011). In the current study no

advantage was found using a segmentation strategy, where our materials required integration of interrelated ideas.

Previously Spanjers et al. (2010) proposed that segmentation reduces transitory effects because it includes both pausing and cueing. Pausing can facilitate learning by providing additional time to process the information. However, if segmentation is explained solely by the pause effect, then a significant segmentation effect should have been found. The lack of an effect in this study suggests that pausing does not automatically produce positive results due to extra processing time of spoken information.

Spanjers et al. (2011) also argued that segmentation could prompt temporal cueing that makes natural boundaries between events in a process or procedure more salient. If cueing occurred in this study it may have had a negative impact by promoting a greater focus on information within the segments and interfered with learners' ability to make connections between segments, leading to a less holistic representation. If cueing occurred and generated negative effects then any advantage gained by pausing was insufficient to produce more positive results overall. Furthermore, according to event boundary research (see Radvansky 2012), segmentation can also promote memory loss under some circumstances. Clearly on tasks that needed to integrate information across segments, a segmentation strategy was ineffective, suggesting that its effectiveness may be dependent upon the nature and content of the tasks and how segmented information is interrelated.

The finding that written text was superior to spoken text supports earlier observations made in communications studies (see Furnham and Gunter 1989; Furnham et al. 1990) and learning English as a foreign language (see Moussa-Inaty et al. 2012). It also supports the multimedia studies of Leahy and Sweller (2011), and Wong et al. (2012), who found that when explanatory text becomes too lengthy, learning is reduced. It is also consistent with research by Singh et al. (2012) who found that for a highly sequential segment task (learning about government bills) reading text was superior to listening to it.

Experiment 2 introduced a multimedia influence by adding a diagram into the presentations. A significant overall multimedia effect was found, indicating that adding a diagram aided both spoken and written text, and thus supports other research into this common phenomenon (see Mayer 2014). A reverse modality effect was also found (spoken text plus diagram was inferior to written text plus diagrams). This reversal of the modality effect (see Low and Sweller 2014) provides further evidence that lengthy spoken text is difficult to deal with due to excessive cognitive load. The advantages of adding spoken text to diagrams or pictures may be lost, or become a disadvantage (as found in this study), if the text is too long. It is notable that the written plus diagram format was presented in a split-attention format as the diagram was separated from the text, which often leads to reduced learning (see Ayres and Sweller 2014). Nevertheless, this presentation format dealt with the transient information in this environment effectively, despite the potential negative effects of split-attention.

Unlike the results found by Singh et al. (2012), no difference was found between spoken and written text, or spoken and spoken segmented text, for the segmented-retention (non-integration) test (Questions 1–8). In fact, no significant differences were found for this test on any of the comparisons. Perhaps crucially the overall success accuracy rates for the segment-retention task were much higher (67% in Experiment 1 and 72% in Experiment 2) than the more complex integration-retention task (54 and 36% respectively). Sweller et al. (2011) have argued that cognitive load theory effects should only occur with high element interactivity materials (Marcus et al. 1996; Sweller and Chandler 1994) where cognitive load is significant. Tests of information contained in segments are fairly low in element interactivity because they are focused on a single concept. In contrast, information that has

to be integrated across segments is by definition much higher in element interactivity because several elements have to be integrated together. This difference suggests that for the segment-retention tasks, spoken text may be less problematical due to low element interactivity. However, it does not explain why Singh et al. (2012) found a difference between written and spoken text on low element interactivity materials.

### Limitations and future research

The absence of an effect on the segment task may also have been influenced by a lack of statistical power. For a 1-way ANOVA with 3 groups, and expected medium/high effect sizes, a total sample of 66–159 participants is required for a power of 80% (calculated using G\*Power©, see Faul et al. 2007). For a  $2 \times 2$  ANOVA (Experiment 2) with expected medium/high effect sizes, a total sample of 52–128 is required. Therefore, the sample sizes for the two experiments must be considered low if medium effect sizes were expected, but more appropriate if high effect sizes were expected. However, the study by Singh et al. (2012) into segmentation and written-spoken text strategies found significant results with large effect sizes, and therefore there was some evidence that large effect sizes could be expected in this study, which was the case. Hence there is some confidence that Type 2 errors did not occur. Nevertheless, future studies should ensure that sample sizes are adequate to detect all significant differences.

The mental effort scores used in this study produced few significant results. As most of the hypotheses generated were based on considerations of cognitive load (see Sweller et al. 2011), this lack of effects was unexpected and may also have occurred due to a lack of statistical power. The single self-rating mental effort measure has been used for over 20 years and has been considered a reliable measure of an index of cognitive load (Paas 1992). More specifically, Paas et al. (2003) consider mental effort to be the actual cognitive capacity allocated to the demands of the task. In terms of when the mental effort scale is best administered, van Gog and Paas (2008) recommend that it should be collected after testing in order to best assess the efficiency of the learning materials. In the present study, it was collected after testing, but produced few significant between-group differences. More recently, it has been proposed that multi-item scales that measure different types of cognitive load (see Leppink et al. 2013), or physiological measures (see Antonenko et al. 2010), are more reliable. Use of such methods may be more fruitful in future research.

In terms of other future research it is assumed that the reason why written text is superior to spoken text is that the former is more permanent, enabling learners to pace the intake of information according to individual reading speeds and revisit information when required. The results of this study support the effect but not necessarily the reasons underpinning it, as these factors were not directly measured. Future research could directly test why written text is superior to spoken text by measuring reading speeds and how many times specific text sections are revisited. Other strategies could also include adding other transient text conditions. For example, the number of segments and how much information (number of words) they include could be varied to further gauge the boundary conditions of such transitory effects.

In terms of generalizing the results, there are a number of factors to consider. The samples used in this study were limited to 15-16 year-old boys. Future research into this area should be conducted with both males and females. Furthermore, future studies should include students of different ages, as well as different learning content. This study, like that of other studies (e.g. Furnham and Gunter 1989; Furnham et al. 1990; Singh et al. 2012) found that written text was superior to spoken text. Hence this result has support in the

literature. However, the finding that written text was superior to segmented spoken text needs more replication, to have confidence that it can be generalized to wide application for integration tasks.

### **Instructional implications**

Spoken information is used extensively in learning environments and is a time-honored method of teaching. Findings in this study have demonstrated that under certain conditions spoken text can hinder learning, arguably as dealing with its transience takes valuable working memory resources away from learning (Leahy and Sweller 2011; Sweller et al. 2011). The consequences that arise from spoken instructional information can be frequently overlooked in an educational setting, thus the transient information effect provides educators with an important instructional design principle. Under many conditions, lengthy explanatory information should be presented in written text to compensate the transience effects associated with spoken text. Learning is likely to be inhibited if educators disregard the transience of spoken text; however these consequences can be avoided by presenting information in a more permanent form. These findings are particularly relevant to online learning environments where audio is a common format used, and these results suggest audio (spoken text) be used with care and with consideration of the length and interactivity of the content.

Depending on the nature of the material students can learn more from either segmented or continuous format. Segmentation can positively affect learning when information needs to be recalled from single spoken segments only (see Singh et al. 2012) or instructional animations (Spanjers et al. 2011). Adding signals or cues can help direct attention towards importation information and task structure (Mayer 2008; van Gog 2014). However, segmenting information into small chunks of information does not always guarantee positive results when information has to be integrated across segments. In contrast, the advantages of permanency using written text compared to spoken text, seems more robust over different types of tasks. Determining which format or medium to use should not be haphazard. A failure to consider transient information effects may result in poorer learning outcomes.

The study investigated three alternatives to spoken explanatory text: continuous written text, segmenting text, and adding a supportive diagram. Overall the results suggest that segmenting spoken text is not such a robust strategy as using written text. On tasks that required information to be integrated across segments, segmentation had a negative impact; whereas written text had clear advantages. Spoken text is of fundamental importance during instructional guidance, however it is evident that learning from spoken text can come at a cost compared to its permanent equivalent (written text). Adding a diagram to text has been found in other studies to be significantly helpful, and in this study adding a diagram to spoken text also lead to significant learning gains compared to spoken text alone (a multimedia effect).

Noting the limitations of the study, three conclusions can be cautiously made from the findings regarding tasks that require the integration of information. Firstly, providing lengthy spoken information without adequate support may be disadvantage due to its lack of permanency. Secondly, substituting written text for spoken text or spoken segmented text leads to higher learning gains. Thirdly, adding a diagram to spoken text enhances learning consistent with the multimedia effect (Mayer 2014).

In a world that is gravitating towards E-learning, the conclusions from this study can be of significance in designing optimal on-line learning environments. Multimedia instruction

has many benefits, but care has to be taken to avoid the negative effects of transitory information caused by lengthy spoken text.

**Acknowledgements** This research was partially supported by an Australian Research Council Grant (DP140103307) to the third author.

## References

- Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, 22, 425–438.
- Ayres, P., & Paas, F. (2007). Can the cognitive load approach make instructional animations more effective? *Applied Cognitive Psychology*, 21, 811–820.
- Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 206–226). New York: Cambridge University Press.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, 8, 47–89.
- Bazermann, C. (1985). Physicists reading physics schema-laden purposes and purpose-laden schema. *Written Communication*, 1, 3–23.
- Byrne, M., & Curtis, R. (2000). Designing health communication: Testing the explanations for the impact of communication medium on effectiveness. *British Journal of Health Psychology*, 5, 189–199.
- Castro-Alonso, J. C., Ayres, P., & Paas, F. (2014). Learning from observing hands in static and animated versions of non-manipulative tasks. *Learning and Instruction*, 34, 11–21.
- Cheon, J., Crooks, S., & Chung, S. (2014). Does segmenting principle counteract modality principle in instructional animation? *British Journal of Educational Technology*, 45, 56–64.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.
- Crooks, S. M., Cheon, J., Inan, F., Ari, F., & Flores, R. (2012). Modality and cueing in multimedia learning: Examining cognitive and perceptual explanations for the modality effect. *Computers in Human Behavior*, 28, 1063–1071.
- Danks, J. H., & End, L. J. (1987). Processing strategies for reading and listening. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 271–294). New York: Academic Press.
- De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, 21(2), 113–140.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191.
- Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 117–134). New York: Cambridge University Press.
- Furnham, A., De Siena, S., & Gunter, B. (2002). Children's and adults' recall of children's news stories in both print and audio-visual presentation modalities. *Applied Cognitive Psychology*, 16, 191–210.
- Furnham, A., & Gunter, B. (1985). Sex, presentation mode and memory for violent and non-violent news. *Journal of Educational Television*, 11, 99–105.
- Furnham, A., & Gunter, B. (1989). The primacy of print: Immediate cued recall of news as a function of the channel of communication. *The Journal of General Psychology*, 116, 305–310.
- Furnham, A., Gunter, B., & Green, A. (1990). Remembering science: The recall of factual information as a function of the presentation mode. *Applied Cognitive Psychology*, 4, 203–212.
- Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, 21, 713–729.
- Hassanabadi, H., Robatjazi, E. S., & Savoji, A. P. (2011). Cognitive consequences of segmentation and modality methods in learning from instructional animations. *Procedia-Social and Behavioral Sciences*, 30, 1481–1487.
- Khacharem, A., Spanjers, I. A., Zoudji, B., Kalyuga, S., & Ripoll, H. (2013). Using segmentation to support the learning from animated soccer scenes: An effect of prior knowledge. *Psychology of Sport and Exercise*, 14, 154–160.
- Kühl, T., Scheiter, K., Gerjets, P., & Gemballa, S. (2011). Can differences in learning strategies explain the benefits of learning from static and dynamic visualizations? *Computers & Education*, 56, 176–187.

- Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, *25*, 943–951.
- Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*, 1058–1072.
- Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 227–246). New York: Cambridge University Press.
- Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*, 49–63.
- Mautone, P. D., & Mayer, R. E. (2001). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, *93*, 377–389.
- Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, *63*, 760–769.
- Mayer, R. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). New York: Cambridge University Press.
- 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., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, *95*, 806–812.
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 277–315). New York: Cambridge University Press.
- Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, *11*, 256–265.
- Mayer, R. E., & Pilegard, C. (2014). 13 Principles for managing essential processing in multimedia Learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316–344). New York: Cambridge University Press.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.
- 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.
- Moussa-Inaty, J., Ayres, P., & Sweller, J. (2012). Improving listening skills in English as a foreign language by reading rather than listening: A cognitive load perspective. *Applied Cognitive Psychology*, *26*, 391–402.
- New South Wales Board of Studies. (2003). *Commerce years 7–10 syllabus*. Sydney: NSW Board of Studies.
- Paas, F. G. (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., 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.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. Oxford: Oxford University Press.
- Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science*, *21*, 269–272.
- Schnotz, W., & Lowe, R. (2008). A unified view of learning from animations and static graphics. In R. Lowe & W. Schnotz (Eds.), *Learning with animations: Research implications for design* (pp. 304–356). New York: Cambridge University Press.
- Schüler, A., Scheiter, K., & Gerjets, P. (2011). Does the modality effect in multimedia learning appear only with text containing spatial information? *Zeitschrift für Pädagogische Psychologie*, *25*, 257–267.
- Schüler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2012). Explaining the modality effect in multimedia learning: Is it due to a lack of temporal contiguity with written text and pictures? *Learning & Instruction*, *22*, 92–102.
- Singh, A., Ayres, P., & Marcus, N. (2012). The transient information effect: Investigating the impact of segmentation on spoken and written text. *Applied Cognitive Psychology*, *26*, 848–853.
- Spanjers, I. A., van Gog, T., & van Merriënboer, J. J. (2010). A theoretical analysis of how segmentation of dynamic visualizations optimizes students' learning. *Educational Psychology Review*, *22*, 411–423.

- Spanjers, I. A., van Gog, T., Wouters, P., & van Merriënboer, J. J. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education, 59*, 274–280.
- Spanjers, I. A., Wouters, P., van Gog, T., & van Merriënboer, J. J. (2011). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior, 27*, 46–52.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185–233.
- Sweller, J., van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.
- Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human Computer Studies, 57*, 247–262.
- van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 263–278). New York: Cambridge University Press.
- van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 16–26.
- Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction, 22*, 449–457.
- Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G. A., Paas, F., et al. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior, 25*, 339–347.
- Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research, 78*, 645–675.