Aligning Affordances of Graphics with Learning Task Requirements

RICHARD LOWE1*, WOLFGANG SCHNOTZ2 and THORSTEN RASCH2

1School of Education, Curtin University, Perth, Australia
2Department of General and Educational Psychology, University of Koblenz-Landau, Landau, Germany

Summary: Dynamic subject matter can be portrayed to learners in the form of either static or dynamic depictions. Two plausible bases for choosing a depiction format are alignment with the subject matter’s dynamics, or the specific affordances a depiction provides for performing a particular learning output task. Experimental participants viewed an ordered set of eight images depicting key stages of a kangaroo hop presented in a dynamic, successive or simultaneous format. Control participants viewed no presentation. The output task required participants to rearrange a random sequence of the eight kangaroo images into the correct order of a kangaroo hop. Those who viewed the successive presentation were most successful in placing the images in their correct order while those in the dynamic condition were least successful. The results contradict the prevailing instructional design orthodoxy that the dynamic properties of a depiction should be aligned with the dynamic nature of the referent content. Copyright © 2010 John Wiley & Sons, Ltd.

Considerable research effort has been devoted to comparing the relative effectiveness of static and animated depictions for supporting learning. However, the results of these comparisons have proven less than conclusive in terms of generalisable effects (e.g. Park & Hopkins, 1992; Tversky, Bauer-Morrison, & Betrancourt, 2002). The facilitative effects of animations on learning appear to occur in specific areas and under specific conditions (Höffler & Leutner, 2007). This raises the possibility that it may be unreasonable to expect either type of depiction to be universally superior to its alternative (c.f. Rey, 2010). Rather, it seems likely that the educational utility of static and animated depictions is closely related to the nature of the learning task they are intended to support (and not the type of depiction per se). However, this appears not to be merely a matter of ensuring that static pictures are used for static content and dynamic pictures are used for dynamic content (Schnotz & Lowe, 2008). Indeed, it is possible that in some cases instructional resources that include static graphics can actually be superior to those containing dynamic alternatives (Mayer, Hegarty, Mayer, & Campbell, 2005). Instead of simply matching depiction type and content, a more sophisticated approach is warranted when choosing which type of depiction may be best in a particular educational resource (c.f. Narayanan & Hegarty, 2002).

Instructional designers need to go beyond an approach that relies on the superficial characteristics of depiction and content to one that gives detailed consideration to the processing activities required for success in an output task. A key question with respect to these activities is: to what extent are the processing affordances (Gibson, 1979) offered by a particular depiction aligned with the type of mental representation that a learner would require in order to cope with a set of task demands? Here, the focus is less on coarse-grained issues such as broad depiction type (i.e. static or dynamic) and more on fine-grained issues that impinge on the construction of a task-appropriate mental representation (Schnotz, 2005).

Static and dynamic external representations clearly differ in the types of information they make available to the learners. These differences can have consequences for the quality of the mental representations that learners are able to construct. A widely held view amongst instructional designers is that dynamic graphics are the obvious choice for depicting dynamic subject matter. It could be argued that this view is warranted because such depictions supply learners with an analogue external representation of dynamic aspects of the content. As a result, dynamic information is available for learners to ‘read off’ and use directly in building their mental representations (Lowe, 2007). In contrast, static depictions of dynamic content require that learners mentally animate the material by themselves without the benefit of an external dynamic reference (Hegarty, 1992). Unless the subject matter is relatively simple and familiar to the learner, this unaided internal generation of dynamics is prone to error.

There is a problem with assuming that dynamic graphics are necessarily beneficial because of their capacity to make information about the content’s dynamics explicitly available to learners. This assumption does not take account of factors such as the characteristics of our perceptual equipment (Wolfe & Horowitz, 2004), the role of working memory in the cognitive processing of temporal change (Ayres, Marcus, Chan, & Qian, 2009), the way we mentally represent dynamic systems (Betrancourt, 2005), or the varied levels of explanation that are relevant to explaining learning with animations (Ainsworth, 2008). The mere availability of information in an external representation does not of itself guarantee its incorporation into a learner’s mental representation of the referent situation. This is because it is first necessary for the learner to successfully extract relevant information from that external representation (Lowe & Boucheix, 2008). Extraction of information from dynamic displays such as animations can be challenging for learners, particularly if the material is complex and unfamiliar (Lowe, 2008). Challenges arise both from the transitory nature of animations (Lowe, 1999) and from the fact that animations are perceived in terms of continuous change rather than

*Correspondence to: Richard Lowe, School of Education, Curtin University, GPO Box U1987, Perth 6845, Western Australia, Australia.
E-mail: r.k.lowe@curtin.edu.au

Copyright © 2010 John Wiley & Sons, Ltd.
discrete images (Schnotz & Lowe, 2008). This continuity is normally considered to be beneficial because it is responsible for the dynamic effect that animation creates. However, the benefit comes at a cost in terms of affordances (Gibson, 1979; Zhang & Patel, 2006) regarding information extraction. The following section canvasses some likely differences in the affordances offered by dynamic and static presentations.

**Dynamic, successive and simultaneous depictions**

The dynamic continuity of an animation relies on a series of images being presented in rapid sequence. This continuity effect can occur even when the images presented comprise no more than a very limited selection of those that would be available in a complete 25 frame per second video recording. Professional animators have found that if they choose the images appropriately, they can create the illusion of continuous dynamics with far fewer frames than are used in videos. The psychological underpinnings of this approach can be traced back to the Phi phenomenon described by Wertheimer (1912) in which as few as two frames can produce a motion illusion. Schnotz and Lowe (2008) note that a set of four frames, when looped, can produce a convincing animation of how a dog runs. When we view this parsimonious animation, it appears dynamically continuous because we can readily infer the intervening temporal changes. However, this illusion is possible only because the chosen four frames depict key stages of the dog’s gait distributed strategically across its locomotion cycle.

The same continuity effect would not be produced if four temporally adjacent, non-key frames were used (such as the first four frames from a 25 fps video). Despite using relatively few frames, parsimonious animations are able to produce a continuity effect provided that they are presented at a sufficiently high frame rate. However, this high frame rate comes at a cost, because each image is available to be viewed only for a fraction of a second, severely limiting a learner’s opportunity to carry out a detailed analysis of the material shown in each frame. For example, there would be insufficient time available to characterise the precise configuration that exists amongst the various graphic forms within each image. A learner could not therefore be expected to extract fine-grained information about the exact changes in configuration that occur from image to image across the animation’s time course. The ‘continuous change effect’ produced by an animation further prejudices extraction of information about configuration changes. This is because the continuity produced by animating a series of images means that viewers tend to follow the individual entities that comprise an image rather than compare them (Schnotz & Lowe, 2008). Different sets of processing affordances and constraints can be produced by manipulating the spatial and temporal arrangements of such groups of frames. Decreasing the frame presentation rate sufficiently turns a dynamic presentation into what effectively becomes a succession of static frames. Reducing the rate to zero and spreading the frames into a row makes their presentation simultaneous. Table 1 summarises the different combinations of affordances and constraints associated with dynamic, successive and simultaneous presentations.

These distinctive affordances and constraints could be expected to have an impact on the types of information that learners preferentially extract. The nature of this extracted information set will in turn constrain the characteristics of the mental representation that a learner is able to construct. If we assume that a learner’s mental representation is the basis for how that individual deals with an output task, this has important consequences for task performance. The information that a learner preferentially extracts because of the particular processing affordances a given presentation type offers will therefore ultimately influence success in the output task.

**Kangaroo locomotion**

The current investigation studied the effect of graphic presentation format on learning about the locomotion of a hopping kangaroo. This example was chosen because research on learning with animation has largely neglected the dynamics of complex biological systems. Further, the kangaroo’s gait is quite unlike more common forms of animal locomotion and therefore sets learners some unique processing challenges. Participants viewed dynamic (animated), successive (static) or simultaneous (static) presentations of a kangaroo hopping cycle and their learning performance was compared with that of a control group who received no presentation. To address the issue of informational equivalence (Larkin and Simon, 1987), the content and viewing duration were standardised. All three presentations used the same eight images (with the same frame size) depicting a kangaroo hop and the same total exposure time. Learning was measured using an output test in which participants arranged these same eight kangaroo hop images (provided to participants as a randomly ordered array) into their correct sequence. Kangaroo locomotion involves complex and hierarchical dynamics. Correct sequencing of

<table>
<thead>
<tr>
<th>Table 1. Presentation characteristics</th>
<th>Dynamic</th>
<th>Successive</th>
<th>Simultaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial arrangement of frames</strong></td>
<td>No separation</td>
<td>No separation</td>
<td>Separated</td>
</tr>
<tr>
<td><strong>Temporal duration of frames</strong></td>
<td>Each frame exposed repeatedly and briefly</td>
<td>Each frame exposed once for an extended period</td>
<td>All frames exposed for the total viewing time</td>
</tr>
<tr>
<td><strong>Affordances</strong></td>
<td>Emphasises overall dynamic continuity</td>
<td>Emphasises local configuration differences</td>
<td>Emphasises global patterns</td>
</tr>
<tr>
<td><strong>Constraints</strong></td>
<td>De-emphasises configuration differences</td>
<td>De-emphasises overall dynamic continuity</td>
<td>De-emphasises configuration differences</td>
</tr>
</tbody>
</table>
its component stages in the output task therefore requires the participant to take account of both macro and micro aspects of the set of images. For the purposes of this study, macro aspects were considered as those involving the kangaroo’s body as a whole while micro aspects involved the kangaroo’s component parts (e.g. legs or tail).

The changes that occur in the kangaroo hopping cycle involve various macro and micro aspects. At the macro level, the hopping cycle includes changes in the overall elevation of the kangaroo’s body and changes in lateral extension as the body is alternately expanded and contracted. At the micro level, the cycle includes local changes in configuration as different body parts perform distinctive movements and changes in separation as the kangaroo’s feet leave then return to the ground. These various macro and micro changes differ in how precisely they indicate the likely position of a particular kangaroo image within the whole hopping sequence. For example, the macro aspect of elevation provides an imprecise indication only of where an image could be positioned. It helps to show whether an image belongs near either end of the sequence (minimum elevation) or near to its middle (maximum elevation). However, on its own, elevation does not provide a foolproof way of locating an image in the first or the second half of the hopping cycle. A similar lack of precision applies with the micro aspect of separation that concerns different distances between the foot and the ground. Separation alone does not indicate which half of the hopping cycle an image comes from. Judging sequence position from extension may be even more difficult because of its small variation and the lack of a useful frame of reference.

In contrast, the micro aspect of configuration provides a very precise way of locating a particular image in its correct position within the sequence (Figure 1). Because each image has a characteristic configuration, orderings that take account of differences in individual configuration patterns are less likely to contain sequencing errors than those based on elevation and separation. Learners who are able to incorporate this type of information into their mental representation are well placed to succeed in an output task that requires image sequencing. However, those who lack this information and so must rely on less precise forms of location guidance only will be ill-equipped for such a task.

Arranging kangaroo images

Given the capacity limits of human information processing, it is proposed that the output task of rearranging the eight randomly ordered kangaroo images into sequence would typically be performed via inter-image comparisons. For example, at a basic level, global comparisons could be used to determine changes in macro level information about the relative elevation of the kangaroo’s body as a whole. In the ascent phase, the next kangaroo in the sequence must be higher than the current one, while in the descent phase, it must be lower. Similarly, local comparison of foot-to-ground distances could be invoked when using the micro aspect of separation as the basis for ordering.

The dynamic presentation was considered least likely to equip learners with the type of mental representation needed for success in the output task. The set of affordances offered by the dynamic presentation favoured extraction of information about elevation changes and separation changes but made it virtually impossible to extract information about configuration changes. Without crucial information about the distinctive differences in image configurations, the mental representations built by learners who viewed the dynamic presentation would be inadequate for the output task. This mental representation should make it relatively easy for them to produce a plausible ascent-descent pattern overall and arrange the images so that the distance of the foot from the ground systematically increased then decreased. However, it would not provide an opportunity for the type of detailed configuration-based checking that is essential to ensure that images are not misplaced in the wrong half of the hopping cycle. On the contrary, learners’ preoccupation with the very bottom of the display (the foot-ground separation area) would likely prevent them from even picking up configuration-based cues from the output task materials themselves.

In contrast, learners viewing the simultaneous and successive presentations would at least have the opportunity to extract information about configuration differences from the display. Extraction of this difference information should be more likely with the successive presentation in which discontinuities between temporally adjacent image pairs are emphasised as each image replaces its predecessor. This contrastive effect is essentially the reverse of the psychological continuity that results from the dynamic presentation. However, extraction of configuration difference information from the simultaneous presentation would be expected to require more deliberate and demanding processing. This is because the images to be compared are separated in space rather than ‘superimposed’ (as they are with the successive presentation).

An analysis based on a simplistic matching of the dynamic characteristics of the presentation to those of the subject matter would predict dynamic presentation of kangaroo hopping to be more beneficial to learners than static depictions. However, an analysis based on the processing requirements of the output task that learners must perform is the foundation for a very different type of prediction. For this reason, it was expected that the distinctive affordances for extracting task relevant information offered by the three presentation conditions would affect learner’s success on the output task. In particular, those in the dynamic condition would be more likely to make more errors in the sequencing of kangaroos images than those in the simultaneous and successive conditions.

Figure 1. Distinctive configuration difference patterns of successive images
METHOD

Design and participants

This study used an independent groups design. Participants were randomly assigned to a control (CON; \( N = 25 \)) condition or one of three experimental conditions. The experimental conditions presented learning materials in one of three ways – (1) dynamic (DYN; \( N = 24 \)), (2) successive static (SUC; \( N = 24 \)) or (3) simultaneous static (SIM; \( N = 25 \)). Ninety-eight Teacher Education students \((M = 21.94 \text{ years}, SD = 5.04)\) from Curtin University, Australia, with no special knowledge of kangaroo locomotion volunteered to take part in the study. Of the participants, 90% were female.

Materials

The learning and output stimuli were presented on a 17 inch LCD computer screen with a 1280 × 1024 pixel display. Each experimental condition featured an 8 second representation of a hopping kangaroo, containing eight images (185 × 130 pixels) of a kangaroo at different stages of its hopping cycle, presented with Microsoft PowerPoint.

The eight images were drawings traced from a video of an actual kangaroo in motion. They were selected from the original full sequence of 25 drawings as the subset that best represented the overall hopping process. Two independent judges chose the eight images to represent the main stages of the hopping cycle.

The DYN condition presented the images continuously for eight cycles with each one presented for 0.125 seconds per cycle, thus creating an animation. The animation was presented in the centre of the screen (Figure 2). The kangaroo therefore appeared to hop on the spot, rather than across the screen. This allowed the hopping mechanism to be observed without interference from also having to track the kangaroo across the display.

The SUC condition presented each of the eight images for 1 second with no time delay between frames (Figure 2). The hopping cycle was presented in the centre of the screen and each subsequent image appeared in the same position, replacing the previous image. The SIM condition presented all eight images together across the screen for 8 seconds (Figure 3).

At the end of each 8 second presentation, a black screen was displayed to prevent the participant from making any further visual inspection of the images.

The output task was computer-based (authored using Adobe Flash\textsuperscript{®} software). Instructions provided on the opening screen informed participants that their task was to arrange a set of pictures into a sequence that correctly reflected an actual kangaroo’s hop.

Participants clicked a start button to commence the output task and were presented with a random arrangement of the eight kangaroo images. The arrangement of kangaroo images differed for each participant. However, none of the sets showed the randomly arranged frames in their correct position in the sequence or next to their correct neighbours.

Procedure

Participants were first given a sequential practice task to make them aware of the output task requirements. After the practice task, participants were informed that they would be required to perform a sequencing task similar to the practice task. The control participants then moved straight to the output task, whereas those in the experimental conditions completed the output task directly after viewing their respective 8 second learning materials.

The instructions for the task directed participants to place the eight kangaroo images into the correct order of a kangaroo hop. The experimental groups were instructed to use information they had gathered from the learning task. The experimenter demonstrated how to perform the task by moving the first image into its correct box and then moving it back to the starting position. The participant then repeated this procedure as confirmation.

Participants proceeded with the task and were allowed to change the position of each image as many times as they wished until satisfied with their arrangement. The movement of each image on the screen was logged by the presentation computer. Once completed, participants pressed a finished button. The final positions were recorded for scoring.

Scoring

Output task scores took account of the number of images in the correct position (P), the number of images in the correct consecutive order (O) and the number of chunks (C; either a run of sequential images or a single image, chunks are underlined in the examples below). For example, a correct sequence would have the images in the order of 1, 2, 3, 4, 5, 6, 7, 8; the score for this sequence was 8 (P) + 8 (O) − 1 (C) = 15. A common erroneous sequence had the images in the following order: 1, 6, 2, 5, 3, 4, 7, 8. The score for this sequence was 3 (P) + 4 (O) − 6 (C) = 1.

Eye-tracking

The distinctive sets of affordances offered by the various forms of presentation could be expected to produce differing visual exploration patterns. Although eye-tracking did not form a central part of this study, information about visual...
exploration could help account for any differences that might be found in output task performance. For this reason, a small subset of participants (N = 12, i.e. 3 per condition) from each group had their eye movements tracked during the presentation and output sequencing task. For these participants, the learning materials were presented on the 17 inch TFT monitor (1280 × 1024 pixels) of a 120 Hz Tobii T120 eye tracker. The Tobii eye tracker is accurate to 0.5°, with a drift of <0.3° and freedom of head movement of 30 × 22 × 30 cm³.

Two areas of interest (AOIs) were created in order to examine patterns of visual exploration during the learning and output tasks. One AOI targeted the lower section of the visual display where separation would be detected (i.e. the region between the feet and the ground). This AOI was labelled bottom. The other AOI covering the remaining section of the visual display was labelled remainder. The AOIs were the same across each of the three experimental conditions and the control condition.

RESULTS

Task completion

There were no significant differences between the conditions for the time taken in seconds (M = 159.06, SD = 91.31) to complete the task (p = .751; f = .11).

Output task scores

Levene’s test for Homogeneity of Variance for the output task score was not violated (p = .34). A one way ANOVA was performed with a main effect found for presentation condition on total output score, F (3, 97) = 4.931, p = .003, η² = .13. The means (and standard deviations) for output task scores in each condition are displayed in Table 2.

Participants in the DYN condition performed significantly worse than those in the SUC condition (Cohen’s d = .98) and those in the CON group¹ (d = .67). Those in the SIM condition also performed significantly worse than SUC group participants (d = .60). Results in the SUC and SIM group were no different than control (d = .46, .18, respectively). Performance by participants in the DYN and SIM groups were no different from one another either (d = .38).

Image arrangement

Figure 4 shows the four main ways that participants tended to arrange the kangaroo images during the output task: (1) with the sequence essentially correct,² (2) with frames belonging to one half of the hopping cycle transposed into the other half, (3) with multiple hops (i.e. feet contacting the ground more than twice) and (4) with a low degree of systemic coherence. These categories were used as the basis for a fine-grained analysis of the output task performance.

<table>
<thead>
<tr>
<th>Condition</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>1.00</td>
<td>3.74</td>
</tr>
<tr>
<td>Dynamic</td>
<td>−1.54</td>
<td>3.85</td>
</tr>
<tr>
<td>Successive</td>
<td>3.17</td>
<td>5.46</td>
</tr>
<tr>
<td>Simultaneous</td>
<td>0.32</td>
<td>3.92</td>
</tr>
</tbody>
</table>

A Chi-square test of goodness-of-fit³ was performed to determine whether the sequencing distributions in each condition matched those of the CON group. The SIM group did not differ significantly from the CON group. However, both the SUC and DYN groups were significantly different from the CON group: χ² (3, N = 24) = 16.81, p < .001; χ² (3, N = 23) = 24.31, p < .001. The frequencies of the arrangement types in each condition are presented in Figure 5.

Participants in the DYN condition were no better at producing a correct arrangement of images than those in the CON group. Further, they were the most likely to produce multiple hop sequences rather than just a single sequence. Those in the SUC condition produced the greatest proportion of correct sequences (more than three times as many as the CON group) and were least likely to transpose images.

Visual interrogation

Data collected from the limited sample whose eye movements were tracked were expected to provide no more than preliminary indications of visual exploration patterns (illustrative only). The following results should therefore be treated with caution. For the 12 participants whose eye movements were tracked, there were different patterns of attention devoted to the two AOIs. These patterns are consistent with what would be expected from the affordances and constraints of the different types of presentation. The percentages of total fixation time spent in the bottom AOI for the three sample participants in each condition (during the learning and output phases) are presented in Table 3.

As expected, the DYN group sample directed the majority of their attentional resources during both phases towards the bottom region (i.e. foot-ground separation zone). During learning, those in the SIM group sample directed the majority of their attention towards the remainder (i.e. upper) regions of the visual display. On the other hand, the SUC group sample distributed their fixations during the learning phase more or less equally between the bottom and remainder regions. During the output task however, their time was spent mostly in the remainder region producing a distribution of attention that was very similar to that of the CON group sample.

DISCUSSION

The results support the prediction that alignment of the presentation format with the dynamic nature of the depicted content would negatively affect performance in the output

¹LSD post-hoc analyses.
²Position and sequence correct except for minor deviations.
³With Yates Correction.
task. Output performance of those in the dynamic group was not only worse than performance of successive participants, but was also worse than that of the controls. These results suggest that the mental representation dynamic participants developed from their exposure to the kangaroo animation actually interfered with their output performance. The effect of this mental representation was apparently so influential that it even prevented dynamic participants from taking proper advantage of the configuration information available to them when manipulating the output task materials.

A detailed breakdown of the successive group’s output scores showed that these participants were far more successful than the controls in producing correct sequences. This is consistent with a capacity to distinguish between images more precisely on the basis of distinctive differences in their configurations because of the affordances offered by the successive presentation. In contrast, the nature of the presentation viewed by the dynamic group favoured the following of entities rather than their comparison. This made it virtually impossible for these participants to access such image-difference information. As a result, they would not have been able to incorporate this form of constraint when constructing their mental representations from the initial presentation.

The limited affordances offered by the dynamic presentation effectively confined participants’ extraction of material to information about (i) differences in elevation of the kangaroo’s body as a whole, and (ii) differences in the separation between the kangaroo’s foot and the ground. As a result, their processing of images in the output task could have been largely dictated by these two imprecise frameworks. Because of the extreme lack of precision in these two types of information, dynamic group participants could easily produce highly error-laden sequences but have no effective means for detecting the errors produced. This explanation is consistent with their lack of sequencing success relative to those in the successive group.

As previously noted, the eye tracking results should be treated with caution. However, the data illustrate differences in exploration patterns. The high proportion of bottom fixations during the output of the dynamic group is consistent with what would be expected from a foot-ground separation based exploration.

The findings from this investigation further challenge the widespread instructional design assumption that the dynamic properties of a graphic presentation should be aligned with

![Figure 4. Correct sequencing and examples of transposed, multiple hop and incoherent sequencing](image)

![Figure 5. Proportion of kangaroo image arrangements by condition](image)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Control (%)</th>
<th>Dynamic (%)</th>
<th>Successive (%)</th>
<th>Simultaneous (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning</td>
<td>54</td>
<td>45</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Output</td>
<td>22</td>
<td>64</td>
<td>33</td>
<td>51</td>
</tr>
</tbody>
</table>

the dynamic character of the depicted subject matter. Instead, it seems more important to focus on the processing requirements of the output task when deciding on depiction type. Even then, the choice is likely not to involve a simple dichotomy between static and dynamic graphics because of the wide range of design options and associated processing affordances available within each of these broad depictive categories. The presentations used in this study exemplified only a small fraction of the many different visuospatial and temporal variations that can be employed in the design of static and dynamic graphics. For example, it is conceivable that a somewhat lower presentation speed in the DYN condition would make it possible to maintain perceptual continuity and yet allow learners to extract configuration information (c.f. Lowe, 2006; Schnott, Lowe, Rasch, Meyer, & Wagadariker, 2007). In this way it may be possible to adjust the affordances of the animation so that they are better aligned with output task requirements. It is possible that this pattern of results could be reversed if the task involved global aspects of locomotion (such as trajectory) rather than configuration as was the case here. This possibility could be investigated in future research. Also, it would be useful to collect eye fixation data from all participants.

A potentially important factor that was not explored in this investigation is the possible effect of top-down contributions to animation processing. For example, had students of Veterinary Science participated rather than Education students, the effect of presentation type may have been less pronounced due to their knowledge of bio-mechanics. However, even those who did participate in this study would have likely recruited some existing relevant domain-general, everyday knowledge about animal locomotion. Future work could investigate the interplay between top-down and bottom-up processing (c.f. Kriz & Hegarty, 2007), particularly with respect to the capacity of background knowledge to compensate for misalignments between affordances and output task requirements. Work comparing meteorologists’ and non-meteorologists’ performance on weather map interpretation tasks shows the importance of domain-specific background knowledge in shaping the way individuals respond to graphics-based tasks (Lowe, 1989, 1993).

CONCLUSION

The findings presented here show that the effects of dynamic presentations on learning are not necessarily benign. Unforeseen negative consequences can follow from the specific mix of affordances that are offered by this type of external representation. This research provides yet another demonstration of how unwise it is for instructional designers to rely solely on their intuitions when making decisions about the use of media in educational resources. Unfortunately, numerous instructional designers still develop multimedia learning environments on an intuitive basis (c.f. Narayanan & Hegarty, 2002). The findings also suggest that the issue of mental representation should be central to making decisions about what types of depiction are most likely to produce particular learning outcomes. The mental representation a learner constructs from a specific depiction can play a key role in determining success with the associated learning output task. Whether that mental representation facilitates or impedes the required task-related processing can be traced back to the affordances that the depiction offered the learner for extraction of task-relevant information. This indicates that instructional designers need to take a two-pronged approach when choosing amongst various depictive possibilities. On one hand, they should conduct a careful analysis of the processing operations that are required for successful performance of the output task. On the other, they should devise depictions whose affordances facilitate learner extraction of information required for building a mental representation that is well suited to those processing demands. In summary, the alignment that really seems to matter in deciding which type of graphic to use is not that between the depiction and the content (e.g. dynamic graphics for dynamic content). Rather, it is the alignment between the depiction’s processing affordances and the output task requirements.

ACKNOWLEDGEMENTS

The authors thank Danielle Beissel, Jamie Lok and Rachael Jones for their fine contributions. Research was supported by a Discovery Grant from the Australian Research Council (DP04511988). We also thank two anonymous reviewers for their helpful comments.

REFERENCES


tific diagrams. Educational Psychology, 9, 27–44.


