The cognitive basis of ONPAR assessment: A white paper

Blake Myers
August 25, 2015

1. Introduction

Over the past few decades, there has been a rising interest in the use of dynamic multimodal environments in learning and assessment. One promising feature of such environments is their potential to provide new avenues by which students can more easily convey their content knowledge and skills. Fulfilling this potential, however, can prove quite difficult. Without careful design, students may be presented with too much or too little information about relevant content, or they may be presented with information through a poorly designed format (e.g., through a cluttered screen containing many extraneous stimuli). Such design features can hinder students in conveying their content knowledge and skills. In a case with too much information, for instance, the task may too easily guide students to a correct response, thereby leaving little room for students to convey their *own* content knowledge and skills. Similarly, in a case of either too little information or poorly presented information, students with relevant content knowledge may be unable to convey their knowledge due to a confusion about which stimuli are relevant to a task or about how such stimuli relate to a target question. Given this, two questions arise as to how dynamic multimodal environments might fulfill their potential to provide new avenues by which students can more easily convey their content knowledge and skills. First, *which* design features (or methods) of dynamic multimodal environments can help fulfill this potential? Second, *how* do such design features help to fulfill this potential?

To address these questions, we consider a novel dynamic multimodal environment for assessing student performance, called ONPAR. ONPAR is designed to assess a wide variety of students; e.g., students with low English language proficiency or other language-related challenges, students with mood and anxiety problems, students with autism, and students without such learning-related challenges (Carr & Kopriva 2013).

What is it about ONPAR that has proven effective in assessing student performance? To answer this, we begin in Section 2 by looking at a number of ONPAR methods that are designed to open up new paths by which students can better convey their content knowledge and skills, and we draw from research in cognitive science to examine how these ONPAR methods succeed in opening up such paths. Then, in Section 3, we consider various concepts that are employed in the videogame literature, and we examine how such concepts can be used to broadly situate ONPAR among other dynamic multimodal environments.

2. Accessing students' content knowledge and skills

As many researchers have noted, meaning is not only grasped through the use of words but also through the use of various other presentation modes (e.g., symbols, shapes, sounds, colors) (Kress 1985, 1986; Kress & van Leeuwen 1996, 2001). Accordingly, data from a number of studies suggest that people acquire better understanding from multimodal environments than from unimodal environments (Mayer 1989; Mayer & Gallini 1990; Clark & Paivio 1991; Mayer & Anderson 1991; Mayer &

Anderson 1992; Mayer et al. 1996; Moreno & Mayer 1999, 2002a; Fletcher & Tobias 2005; Low & Sweller 2005; Moreno 2006; Moreno & Mayer 2007; Mayer 2009).

In light of such work, ONPAR tasks employ many types of stimuli, including colors, objects, symbols, numbers, words, written information, spoken information, and spatial layout. For instance, an assessment task on gas exchange includes animations with objects (e.g., test tubes, plants, animals, liquid droppers, containers of oxygen, and containers of carbon dioxide), colors (e.g., blue water, green solution, yellow lights, and green plants), and optional written and spoken questions (e.g., “What color will the water be in each test tube in light and dark?”).

While the diversity of presentation modes is a key component of ONPAR assessment, it is more precisely the *manner* in which the presentation modes are used that is crucial to ONPAR. Broadly described, ONPAR methods are designed to provide reliable entry into student memory, and thus, reliable access to students’ content knowledge and skills. To illustrate, consider two memory systems that play a role in student assessment: long-term memory and working memory. Long-term memory is a system with virtually limitless storage in which information can be held offline for long periods of time and then later retrieved for use. Working memory, by contrast, is a limited-capacity system that is responsible for the conscious processing and manipulation of just a few pieces of information received from long-term memory or sensory systems (Baddeley & Della Sala 1996; Gobet & Simon 1996; Paas 2003; Dehn 2008; Sweller et al. 2011).

The connections between long-term memory and working memory are bidirectional. That is, working memory not only receives information from long-term memory, but it also sends information to long-term memory for later retrieval. This occurs most efficiently when one can actively connect information in working memory to information that already exists in long-term memory. Similarly, information in long-term memory is most efficiently transferred to working memory when the information in working memory shares certain features with the information in long-term memory (Baddeley & Logie 1999; Dehn 2008; Sweller et al. 2011).

For the purposes of educational assessment, three points are particularly relevant. First, information in long-term memory is not merely stored as a list of discrete facts and procedures, but rather as an interrelated network of associated concepts, propositions, strategies, and procedures (Dehn 2008; Sweller et al. 2011). Thus, if one wants to assess students' content knowledge and skills, it is worth asking not only whether the student knows certain stand-alone facts but also how such facts are related to other concepts, propositions, strategies, and procedures in student memory (Pellegrino et al. 2001). One way to access such information is through the use of open-ended questions. Such questions not only provide students with an avenue to more fully portray their content knowledge and skills, but they also help rule out alternative explanations as to how a student arrived at a correct answer. For instance, it is more likely that a student who correctly responds to a true-false question does so by guessing than it is that a student who correctly responds to a short-answer question does so by guessing. When using open-ended questions, however, it is important for the questions to be worded and contextually situated in such a way that students who have relevant content knowledge can discern what the questions are asking. For instance, if contextual information is poorly presented or the question is too open-ended, students with relevant content knowledge might be misled about what the question is asking or about which stimuli are relevant to the question. To help ensure that incorrect answers result from a lack of students' content knowledge, and not from a failure of the test to convey relevant information, ONPAR tends to use what we call "assisted open-ended questions". For example, in a biology task, students are presented with the question, "Why does the water remain green?", along with several nouns (e.g., "light", "energy", "oxygen", and "animal") and verb phrases (e.g., "released by"). Then, students are required to drag and drop the words and phrases to construct a correct response to the question. Similar response-related activities elicited by ONPAR tasks include dragging and dropping minerals in response to the instruction "Order the minerals from soft to hard", positioning objects at various depths in a container of water with respect to their mass and volume, and dragging and dropping bacteria and connector lines in response to the instruction "Use the amino acid differences to make a cladogram for the bacteria". This use of combining fairly open-ended questions or instructions with an organizing structure of relevant response items allows students to convey their content knowledge and skills without being misled about what a question is asking or about which stimuli are most relevant to a task.

The second point about how ONPAR methods provide reliable access to students' content knowledge and skills pertains to the use of schemas in learning and assessment. Roughly, schemas are organizational structures in long-term memory that serve to encode students' knowledge and skills in an interrelated network of concepts, propositions, strategies, and procedures (Dehn 2008). Although working memory can only process a few pieces of information at a given time, existing schemas in long-term memory can be used to group together or “chunk” (Miller 1956) several pieces of information into a single informational structure. This informational structure can then be efficiently processed as a single piece of information, thereby avoiding the strain on working memory that might otherwise result from processing each piece of information in isolation (Sweller et al. 2011; Guida et al. 2012). This has clear relevance to student assessment. Consider, for instance, an ONPAR task on food chains. Students are initially presented with two illustrations. The first illustration, labeled “Food Chain”, depicts a series of images interpolated with three rightward arrows, portrayed in the following order: [image of the sun], [rightward arrow], [image of a plant], [rightward arrow], [image of a rabbit], [rightward arrow], [image of a fox]. The second illustration, which is positioned directly below the first and labeled “Number of Living Things”, is a bar graph representing the population size of the three organisms from the food chain: the plant (which has an initial population of 12), the rabbit (which has an initial population of 8), and the fox (which has an initial population of 4). After observing the illustrations, students can then press the “GO” button to move to the second stage of the task. In this stage, students watch a short animation in which a red “X” is placed over the image of the rabbit in the food chain illustration, and subsequently, the graph bar representing the rabbit population decreases from 8 to 0. Students must then estimate how the graph bars representing the plant and fox populations should change in light of the changes in the rabbit population.

With this example in mind, consider how a student’s knowledge schema might facilitate the student’s ability to identify and process relevant information. Suppose, for instance, that a student possesses a schema of food chains and population size;

particularly, one that serves to group organisms together in accordance with a linear sequence of production and consumption (where an increase in the population of one organism tends to lead to a decrease in the population of organisms lower on the food chain and an increase in the population of organisms higher on the food chain). Such schemas can prime students to quickly detect and attend to relevant stimuli; e.g., by priming them to quickly categorize stimuli as either organisms or relations between organisms. By contrast, students without such a schema are more likely to perceive various stimuli (e.g., arrows, images, bars) in isolation from one another, and thus are less likely to discern the meaning or relevance of the stimuli, since the meaning and relevance are largely determined by how the stimuli relate to one another.

Given this, consider how ONPAR's multimodal design serves to differentiate students with relevant content knowledge from those without it. Compare, for example, ONPAR design with that of more traditional text-based assessments. In text-based assessments, students with language-related challenges may be unable to convey their relevant content knowledge and skills; for, an assessment environment that uses only text-based stimuli is likely insufficient for activating existing schemas in students with language-related challenges. This can be problematic, since without the activation of students' relevant knowledge schemas, students are likely to have difficulty in identifying and processing relevant information (Sweller et al. 2011), and thus are likely to perform as if they lacked the relevant content knowledge that they in fact have.

ONPAR tasks aim to avoid this problem by providing multiple paths by which students' relevant knowledge schemas might be activated. To illustrate, consider again the food chain task. Throughout the task, students are given information through multiple presentation modes (e.g., images, shapes, symbols, colors, spatial layout, rollover icons, and written and spoken text). Thus, if a student possesses relevant content knowledge, and thus a relevant knowledge schema, then although certain presentation modes (e.g., written text) might fail to convey information in a way that activates the student's relevant knowledge schema, one of the other many presentation modes may well still succeed in doing so, and thus may well facilitate the student's ability to identify and process relevant information. Also, this use of multiple presentation modes in no way "dumbs down" the task, since any student who lacks the relevant content knowledge also lacks a relevant knowledge schema; and without the activation of a relevant knowledge schema, students remain at a significant disadvantage in identifying and processing relevant information, and thus in providing a correct response.

The third point about how ONPAR methods provide reliable entry into student memory is closely related to the second point. However, while the second point focuses more broadly on the role of multiple presentation modes in activating students' knowledge schemas, the third point focuses more closely on how specific design features of ONPAR (such as the particular types of stimuli used and the particular organization of such stimuli) serve to facilitate efficient processing of information in students with relevant content knowledge and skills. To illustrate, during knowledge-based assessments, students must transfer information from long-term memory to working memory, where they can then consciously processes it in order to provide a response (Mayer 2009; Sweller et al. 2011). However, given the limited capacity of working memory, students who have relevant content knowledge may be unable to transfer it from long-term memory to working memory under certain conditions, and thus may fail to provide a correct response. For example, if students are assessed in an environment in which information is presented on a cluttered screen, it may appear as if students with relevant content knowledge fail to have it. The reason is that by requiring extraneous processing in working memory (which might result from the student's need to discern relevant information amongst a background of clutter), working memory may temporarily lack the resources to retrieve relevant information from long-term memory, and thus, students with relevant content knowledge might be unable to use such information to provide a correct response. As Baker (2009) notes with respect to verbal stimuli:[1]

> Difficulty can inhere in structural and usage aspects of verbal stimulus materials to a degree that often swamps content knowledge.… It is important, then, that difficulty of assessments and the inferences made from lower performance is attributable to performance in the task domain of interest and is not confounded by language syntax that pose additional hurdles for the examinee.

A similar point can be made with respect to non-verbal stimuli: namely, difficulties that are unrelated to the task domain of interest may exist in structural and usage aspects of non-verbal stimuli to a degree that swamps content knowledge. While it is clear that such interference should be removed from assessment tasks, doing so requires the ability to identify the interferences, which can prove quite challenging when working with dynamic multimodal environments such as ONPAR. Research in cognitive science, therefore, is particularly useful here, since it reveals a wide range of features that may unexpectedly

---

[1] Also see Cole & Moss 1993; Pellegrino et al. 2001.

interfere with the retrieval of information from long-term memory and thereby interfere with students' ability to convey their content knowledge and skills.

Several studies, for instance, have found that when words are not presented in close spatial proximity to corresponding illustrations, people show decreased understanding of relevant content material (Mayer 1989; Sweller et al. 1990; Chandler & Sweller 1991; Mayer et al. 1995; Tindall-Ford et al. 1997; Moreno & Mayer 1999). Mayer (1989) presented a group of students with an illustration on the workings of car brakes, along with various written descriptions placed next to the relevant parts of the illustration. A second group of students was presented with the same illustrations and written descriptions, but the illustrations were placed at the top of the page, while the corresponding written descriptions were placed at the bottom of the page. After viewing the material, students were asked a series of transfer questions about the workings of car brakes (e.g., "What could be done to make brakes more reliable, that is, to make sure they would not fail?"). Students in the second group correctly answered fewer transfer questions than those in the first group. Similar results were found with respect to the positions of text and symbols in geometry diagrams (i.e., those who received booklets with the text and symbols positioned a fair distance below the diagram took significantly longer to solve transfer problems than those who received booklets with the text and symbols placed in close proximity to the corresponding parts of the diagram).

While such research is largely aimed at discovering how people *learn* in multimodal environments, it also has clear applications to multimodal *assessment*. In ONPAR tasks, interactions between the students and the environment do not merely involve a series of questions and responses. Rather, prior to presenting the student with any questions, ONPAR tasks first present the student with a diverse range of stimuli that are designed to orient the student towards relevant information and thereby prepare the student for a target question. Thus, as students move throughout a task, they must learn to identify and assign appropriate meanings to relevant stimuli. Given this, learning-related research (e.g., that on the relationship between learning and spatial proximity) is often germane to the design of multimodal assessments. To illustrate, if ONPAR tasks were designed so that written texts were positioned in far spatial proximity from corresponding illustrations, students may have greater difficulty learning to identify and assign appropriate meanings to relevant stimuli, and hence may have greater difficulty in conveying their content knowledge and skills.

In light of the aforementioned research on spatial proximity in learning, ONPAR tasks present written text in close proximity to corresponding illustrations. In a chemistry task, for example, illustrations of four different molecular models are presented in close proximity to one another and positioned in the center of the screen. Moreover, the description for each model (e.g., "Space-filling model", "Lewis structure", "Ball-and-stick model") is presented directly beneath the corresponding illustration.

In addition to the way in which ONPAR uses spatial *proximity* in its design, it also ONPAR uses spatial *consistency* to facilitate student interaction with onscreen stimuli. Research suggests that the visual system is highly selective to spatial regularities in the environment (Jiang et al. 2013). Thus, to better enable students to use onscreen support tools in ONPAR tasks, the location of such tools remains constant throughout tasks. For example, navigational buttons are always located at the bottom center of the screen and remain in the same order, target questions are always located near the top left of the screen and are positioned next to a button that can be used to hear a spoken version of the question, and rollover icons (which can be used to view a demonstration of how one is to perform a given action) are always located at the top right of the screen.

As is the case with spatial features, temporal features have the potential to either clear or obstruct paths by which students might convey their content knowledge and skills. Several studies suggest that without close temporal proximity between the presentation of a spoken text and that of a corresponding animation, students must hold the spoken information in working memory while waiting to identify the corresponding animation. This wastes mental resources. Spoken information should be presented at the same time as the corresponding animation (Mayer and Anderson 1991, 1992; Mayer and Sims 1994; Mayer et al. 1999). In a study by Mayer et al. (1999), students were presented with an animation about either lightning formation or the workings or car brakes, as well as a corresponding spoken narration about the animation's content. Students were divided into three groups. The first group watched all of the animation while concurrently listening to all of the spoken narration. The second group listened to short segments of the spoken narration that were interpolated between short segments of the animation. The third group listened to all of the spoken narration either before or after watching all of the animation. Students were assessed on the basis of retention, transfer, and matching. Retention was measured by giving students 6 minutes to write down everything they could remember about lightning formation (or the workings of car brakes); those who provided more of the explanatory steps that were portrayed in the animation were judged to have better retention. Transfer was measured by giving students a series of questions to answer (e.g., "Why would clouds appear without lightning?”); those who provided a higher number of correct answers to the questions were judged to have better transfer ability. Finally, matching was measured by asking students to label parts of a graphic illustration (e.g., students might be asked to “circle part of the brake line and write L next to it”); those who correctly labeled more of the relevant parts were judged to have better matching abilities. On all three measures (retention, transfer, and matching), students in the third group (i.e., those who listened to all of the spoken narration either before or after watching all of the animation) showed decreased performance when compared to the first and second groups. Similar results were found by Mayer and Anderson (1991, 1992), and Mayer and Sims (1994).

Given such results, ONPAR provides students with (optional) written and spoken stimuli that are presented at the same time as any corresponding non-verbal stimuli. Moreover, throughout ONPAR tasks, students can listen to or replay spoken information by clicking on the “ENGLISH” (or “TRANSLATE”) button positioned next to the written text. The written text, as well as the button used to listen to and replay the spoken version of the text, remains on the screen throughout the corresponding animation. A mathematics task, for instance, involves an animation in which the numeral “1” is passed through a machine that is labeled “+3”, and the numeral “4” comes out from the other side of the machine. During the animation, the sentence “This is a number machine” is written at the top of the screen. The “ENGLISH” (or “TRANSLATE”) button positioned next to the sentence can be pressed before, during, or after the animation to hear the sentence read aloud. This ensures that students do not have to hold spoken information in working memory while waiting to identify the corresponding animation.

A further feature that can affect students' ability to convey their content knowledge and skills pertains to the way in which verbal stimuli are presented. Research suggests that people perform worse when language is presented in a formal style rather than a conversational style, such as when third-person constructions are used rather than first- or second-person constructions (e.g., when using "people" in place of "you", or "the" in place of "your") (Moreno & Mayer 2000, 2004; Mayer et al. 2004). In a study by Moreno and Mayer (2000), students watched an animation on lightning formation while listening to a corresponding narration that described the main steps of lightning formation. Some students received personalized narration in which the speaker addressed students directly and used first- and second-person constructions (such as, "I" and "you"), while other students received nonpersonalized narration in which the speaker neither addressed students directly nor used first- or second-person constructions. The key content of the narration was the same for both groups (i.e., describing the same key steps of lightning formation). Those in the nonpersonalized group exhibited decreased performance on transfer tests when compared to those in the personalized group. Transfer tests consisted of solving problems about lightning formation on the basis of what was learned in the presentation (e.g., students might be asked, "What does air temperature have to do with lightning?").

Similar effects have been found with other subject areas, such as the human respiratory system (Mayer et al. 2004) and environmental science (Moreno & Mayer 2000, 2004). In a study by Mayer et al. (2004), students were presented with a short animation along with a corresponding spoken narration that described the workings of the human respiratory system. One group of students (the nonpersonalized group) received a spoken narration with the word "the" used in several places, while another group of students (the personalized group) received the same spoken narration except that twelve occurrences of the word "the" were replaced with "your". Those in the nonpersonalized showed decreased performance on transfer tests when compared to those in the personalized group.

In light of such findings, ONPAR tasks often word questions, instructions, and descriptions in a conversational style. In a mathematics task, for instance, instructions are given in a second-person construction; e.g., "Show how you got your answer". The task also uses language that refers to personalized situations; e.g., "Emma wants a game player and games." Such phrasing serves to prevent interference that might otherwise occur from the use formal (rather than conversational) language.

A second issue related to the personalization of language concerns the type of voice that is used in spoken narrations. In a study by Atkinson et al. (2005), students were presented with worked-out word problems in arithmetic along with an onscreen pedagogical agent. For some subjects, the pedagogical agent spoke in a natural human voice, while for other subjects, the agent spoke in a computerized voice. Those who were presented with the computerized voice performed worse on both near and far transfer problems than those who were presented with the human voice. Near transfer was measured with word problems that were structurally similar to previous example problems but with a different surface story line. Far transfer was measured with word problems that differed in both the deeper structure of the problem and the surface story line.

Similarly, in a study by Mayer, Sobko, and Mautone (2003), students were presented with an animation on lightning formation along with a corresponding spoken narration describing the steps of how lightning forms. One set of students heard the narration in a foreign-accented voice, while another set of students heard the narration in a standard-accented voice. Those who heard the narration in a foreign-accented voice performed significantly worse on transfer problems than those who heard the narration in a standard-accented voice.

In accordance with these results, ONPAR tasks present students with (optional) spoken information in a standard-accented human voice. In a physics task, for example, students are instructed, “Watch what happens. Describe the forces on the box.” The instructions are spoken in a female human voice with a standard accent.

Another feature that can impact students’ ability to convey their content knowledge and skills pertains to the use of interesting yet extraneous material. Harp & Mayer (1997) presented students with a booklet on lightning formation. One set of students received a condensed version that described the steps of lightning formation in five paragraphs, with each paragraph paired with a corresponding illustration. Another set of students received an extended version of the booklet that contained everything from the condensed version, along with five additional photos (one placed next to each paragraph) and additional written text that conveyed interesting facts about the photos but were irrelevant to the steps of lightning formation. For example, a photo of a golfer was paired with the sentence, “Golfers are prime targets of lightning strikes because they tend to stand in open grassy fields, or to huddle under trees.” After reading the booklet, students were given a series of transfer problems; e.g., the question “What does air temperature have to do with lightning?" Students who received the extended booklet performed significantly worse on transfer problems than those who received the condensed version. Similar results were found in a computer-based study (Mayer et al. 2001) and in other booklet-based studies (Harp & Mayer 1998).

While further investigation is needed to determine whether the added material reduces understanding because it is extraneous or because it is interesting, there is reason to think that both features play a role. Mayer et al. (1996) presented students with a written overview of ocean wave formation. One set of students received a relatively brief account of ocean wave formation, while another set received all of the material from the brief account along with extraneous information (i.e., information that was not incorporated into transfer questions) about quantifying ocean wave formation. Students who received the extended version performed significantly worse on transfer problems than students who received the brief version. Since it seems unlikely that the information on quantifying ocean wave formation was particularly interesting to students (at least relative to the other information), the presentation of extraneous material (regardless of whether it is interesting) seems sufficient to reduce student understanding.

However, this is not to say that the interestingness of extraneous material has no effect on student understanding. For, it may well exacerbate the drop in student understanding by more pronouncedly directing students' attention to irrelevant features. In support of this, Garner et al. (1991) found that students remember interesting yet extraneous material better than they remember the key material.

Given such results, ONPAR tasks only use images, animations, and other stimuli that are directly relevant to the key content. Moreover, any text that is used is optional and limited to basic instruction. For example, in a mathematics task, students are shown that 2 blue squares added together equals 6. This is done by displaying 2 blue squares on the screen with an addition sign between them, along with an equals sign positioned to the right of the rightmost blue square, and the numeral "6" positioned to the right of the equals sign. There is no additional extraneous material added to the display that could hinder students in retrieving relevant information from long-term memory.

Due to the vast array of potentially relevant stimuli in dynamic multimodal environments, there are often cases in which it is useful to employ a more direct method of guiding students towards relevant stimuli; namely, through the use of signals. Signals are added to guide students' attention to the relevant material. They do not provide the student with new information. They merely "highlight (or repeat) the essential material in the lesson and … guide learners' organization of the essential material into a coherent structure" (Mayer 2009). In a study by Atkinson (2002), subjects were presented with multistep example word problems on mathematical proportions. Some subjects were presented with example problems along with spoken explanations of the steps to solve the problem. Other subjects were presented with the same example problems and the same spoken explanations, but they were also presented with an onscreen agent that directed subjects' attention (via gesture and gaze) towards the relevant step that was being explained. Atkinson (2002) found that subjects who were presented with the onscreen agent performed better on both near and far transfer tests than subjects who only received the example problem and spoken explanations. Near transfer was measured by having subjects solve problems that were dissimilar in storyline but similar in structure to the example problems, while far transfer was measured by having subjects solve problems that were dissimilar in both storyline and structure to the example problems.

Similarly, Jeung et al. (1997) presented two groups of subjects with the same example geometry problems, which were accompanied by the same spoken explanations of the problems. However, for just one of the groups, the part of the problem that was being described in the spoken explanation would flash on the screen. Jeung et al. (1997) found that in visually complex tasks, those who were presented with the onscreen flashes performed better on problem-solving posttests than those who did not receive the onscreen flashes. Research also suggests that people benefit from a variety of other types of cues, such as graphic displays of the structural organization of a written passage (Stull & Mayer 2007), vocal emphasis (Mautone & Mayer 2001), and organizational headings and numbers within a text (Harp & Mayer 1998; Mautone & Mayer 2001).

In accordance with such research, ONPAR tasks include a number of design features that help direct students' attention to relevant information; e.g., through the movement and position of objects on the screen, the color of objects, changes in the color of objects, and the use of (optional) written and spoken instructions. In certain ONPAR tasks, for example, the key words and phrases used in written questions and instructions are highlighted in order to direct students' attention to such words and phrases. To illustrate, in a mathematics task, students are asked, "How much does 1 ball weigh?" In the written version of the question, the phrase "1 ball" and the term "weigh" are both colored blue and underlined. When the cursor is dragged over the phrase "1 ball", 1 ball is outlined in blue. When the cursor is dragged over the term "weigh", the scales are outlined in blue.

While many of ONPAR's design features are constructed in accordance with findings in cognitive science, the success of ONPAR also points to new areas of research that may have promise. For instance, consider the finding that students' ability to convey their content knowledge and skills can be affected by the manner in which graphics are presented with verbal stimuli. Particularly, such studies show that students have more difficulty learning material when graphics are presented with written text than when they are presented with spoken text (Jeung et al. 1997; Mayer and Moreno 1998; Kalyuga et al. 1999, 2000; O'Neil et al. 2000; Moreno et al. 2001; Moreno & Mayer 2002). In a study by Mayer and Moreno (1998), students were presented with an animation on either lightning formation or the workings of car brakes. All students also received descriptions about the key steps of lightning formation or the workings of car brakes. But while some students received the descriptions via spoken text, other students received them via written text. Mayer and Moreno (1998) found that in both cases (i.e., that of lightning formation and that of car brakes), students who were presented with both an animation and written text exhibited decreased performance when compared with students who were presented with both an animation and spoken text. Specifically, those in the written text group showed decreased performed on tasks in which students were asked to write down an explanation of either lightning formation or the workings of car brakes. They also showed decreased performance on tasks in which students were asked to find and label elements from the animation (e.g., finding the updraft and writing U next to it) and on transfer problems (e.g., answering questions such as, "What could you do to decrease the intensity of lightning?"). Similar effects have been found across a diverse range of tasks and topics, including mathematics problems (Jeung et al. 1997), environmental science simulations (Moreno et al. 2001; Moreno & Mayer 2002b), aircraft simulations (O'Neil et al. 2000), the workings of electric motors (Mayer et al. 2003), and tasks in which diagrams were used in place of animations (Jeung et al. 1997; Kalyuga et al. 1999, 2000).

How do such findings pertain to ONPAR design? While ONPAR tasks often present students with graphics, spoken text, and written text, students are able to *choose* the specific presentation modes with which they work. In a mathematics task, for instance, there are a variety of ways in which students might learn that two scales are balanced; e.g., students might read the sentence "The scales are balanced" at the top left of the screen, or they might click the button labeled "ENGLISH" or the button labeled "TRANSLATE" to listen to a spoken version of the sentence in either English or another language, or they might observe the onscreen graphics to see that both sides of the scale are at the same height. While it seems plausible that this type of autonomy to choose the presentation modes with which one works might reduce the threat of interference from the combined presentation of graphics, spoken text, and written text, further experimentation is needed to explore this possibility.

The role of student autonomy in dynamic multimodal environments also raises a more general point about the diversity of learning and testing styles. As many have noted, not all people work well with the same learning or testing style (Felder & Silverman 1988; Jonassen & Grabowski 1993; Ford & Chen 2001; Bajraktarevic et al. 2003; Graf & Kinshuck 2008). Videogames respond to this issue either by having a game design that allows for diverse styles of play or by allowing players to choose the settings to best fit their style of play (Gee 2007). Similarly, ONPAR tasks give students a significant amount of autonomy over their environments, such as control over the presentation modes with which they interact and control over the pacing of tasks.

Accordingly, research in cognitive science suggests that multimodal environments can be effective in assessing students across a wide variety of learning and testing styles. Many students who are considered to be at a disadvantage show heightened performance when learning or testing in multimodal environments. For example, those with low spatial ability show enhanced performance when animations are used as opposed to still images or no images (Hays 1996; Lohr & Gall 2008); and while those with low working memory capacity tend to struggle with instruction and assessment that contains only written text, they benefit when written text and images are used together (Wey & Waugh 1993; Graf & Kinshuck 2008).

Given that students vary in their cognitive styles and abilities, it is often important to integrate different learning and testing styles into education, since doing so can ease learning and allow for more accurate assessment (Jonassen & Grabowski 1993; Graf & Kinshuck 2008). Not surprisingly, students who are disposed to a particular learning or testing style often have difficulty with assessments that do not match that learning or testing style (Felder & Silverman 1988; Ford & Chen 2001; Bajraktarevic et al. 2003; Graf & Kinshuck 2008). Moreover, the better the fit between a student's testing style and the presentation form of the material, the more likely it is that the student's responses will accurately reflect the student's content knowledge and skills (Graf & Kinshuck 2008).

In light of the large variety of learning and testing styles, such findings may initially seem to suggest that the more presentation modes there are, the more likely it is that the material will facilitate a wider range of students in conveying their content knowledge and skills. However, this is not always the case. Several studies have found that people achieve better understanding when the same information is not given in multiple formats; e.g., they achieve better understanding when graphics and spoken text are presented alone than when they are presented with written text (Mousavi et al. 1995; Kalyuga et al. 1999, 2000; Craig et al. 2002; Mayer et al. 2001; Moreno & Mayer 2002b; Sweller 2005).

Mousavi et al. (1995) presented one group of students with worked-out geometry examples along with diagrams, written text, and a spoken version of the text. Another group of students was presented with the same geometry examples, diagrams, and written text but without any spoken text. Students in the former group (i.e., those who received linguistic information in both a spoken and written form) showed decreased performance on transfer tests relative to students in the latter group (i.e., those who received linguistic information only in written form). Transfer tests required students to use the same geometric theorems that were used in the worked-out examples but in a context with different diagrams.

Similar effects have been found with other topics, such as electrical engineering (Kalyuga et al. 1999, 2000), lightning formation (Craig et al. 2002; Mayer et al. 2001; Moreno & Mayer 2002a), and environmental science (Moreno & Mayer 2002b). Mayer proposes that this effect is due to the fact that redundant information (from presenting the same linguistic information in multiple forms) increases cognitive load and thereby decreases performance (Mayer 1997, 2009; Kalyuga et al. 1999). This suggests a possible tension between designing assessments that cater to a wide variety of learning and testing styles (i.e., by using multiple presentation modes) and designing assessments that do not unintentionally hinder students' performance by increasing cognitive load.

As noted above, ONPAR responds to this type of issue in much the same way as good videogames do: i.e., by offering not only multiple, but also *optional*, presentation modes. While ONPAR tasks often contain the same, or similar, information in different formats, students are able to choose the formats through which they interact and receive information. For instance, in a mathematics task, students are presented with a problem in which the numeral "5" is inserted into a machine that is labeled "+3", and a question mark comes out of the machine. At this point, students can attempt to solve the problem by inserting a numeral into the box with a question mark. Or for further clarification, they can read the corresponding text that says, "The rule is +3. What number comes out?", or they can listen to a spoken version of the text. While this use of multiple yet optional presentation modes may help explain why a wide variety of students show increased performance on ONPAR tasks (Carr & Kopriva 2013), further research is needed to understand the precise role of student autonomy in multimodal environments, and in particular, how such autonomy might be used to increase the number of presentation modes without unintentionally increasing cognitive load.

3. Situating ONPAR among other dynamic multimodal environments

Given the promise of ONPAR in opening up paths by which students can better convey their content knowledge and skills, consider more broadly the relationship between ONPAR and other dynamic multimodal environments, most notably videogames. While much of the relevant literature on videogames is more conceptual than experimental, it is nonetheless valuable when considering multimodal environments. As McClarty et al. (2012) point out:

> Most of the available studies consist of descriptive analysis of the impact games have on students' attitude towards the subject being taught and their motivation to attend and engage in class.... In rare occasions when researchers have attempted to investigate the relationship between learning within digital games and academic performance, the results are mixed because of differences in definitions and methodologies.... [C]reating definitions and models for many of the attributes that are considered integral parts of the power of games (e.g., motivation, engagement, agency) would ... allow for a more coherent research approach.

Thus, in order to determine the importance of various game-like features in learning and assessment, relevant concepts are needed to pick out such features and differentiate one feature from another. Accordingly, for the purposes of this paper, such conceptual work is useful in its ability to broadly situate ONPAR among other dynamic multimodal environments.

Consider, for instance, a distinction between two types of educational games: *conceptually embedded games* and *conceptually integrated games*. In conceptually embedded games, the player explores the game world and gains scientific understanding through overt inquiry at particular sites (Gee 2007; Squire 2012). In conceptually integrated games, by contrast, the relevant science concepts are "integrated directly into the core mechanics that operate in the game environment" (Clark & Martinez-Garza 2012). Conceptually embedded games tend to make it easier to teach and assess key ideas in an explicit (as opposed to tacit) way, while conceptually integrated games often make it more difficult for players to explicitly articulate their content knowledge (Masson et al. 2010).

ONPAR tasks share a number of key features with conceptually embedded games. Throughout a given task, students must engage in overt inquiry about scientific or mathematical content (e.g., about the relationship between temperature and reaction time, or about what happens when a ball with a particular density and volume is dropped into a liquid). As in conceptually embedded games, such inquiry takes place in an environment in which target questions are situated within a broader framework of relevant contextual features. For instance, ONPAR tasks typically begin by presenting students with an introductory context that situates them within an environment that is closely tied to the target question. Subsequently, students are presented with further stimuli that build upon the introductory context to better orient the students towards the target question and thus provide them with a clear avenue through which they can convey explicit content-related knowledge. This transition (from the introductory context, to the buildup stimuli, to the target question) serves to restrict information to just a few key features at the beginning of the task, and then to increase the amount of relevant stimuli as the student moves towards the target question. This is important, since if early interactions between the student and environment are too complicated or too open-ended, students may be prone to form hypotheses that work in a given situation but have little extension to later more complex interactions. As many have noted, earlier interactions should be less complex, involve fewer pieces, and be easier to complete (Clark 1989; Elman 1991a, b; Gee 1992, 2001; Steinkuehler 2013).

In support of this, Kersten and Earles (2001) presented subjects with an artificial language and found that those who were first presented with single words and only later presented with entire sentences learned the meanings and morphology of the words better than those who were only presented with entire sentences. Similar results have been found in research on neural networks. Networks that begin with a restricted amount of data along with limited working memory (or limited access to prior internal states) perform better in learning artificial languages and are less likely to make false generalizations than networks that begin with more data and better access to prior internal states (Clark 1989; Elman 1991a, b, 1993).

Accordingly, ONPAR tasks typically begin by presenting students with an animation or illustration. Students must then make inferences about the relevance of the initial stimuli; e.g., the color of the water, the labels “O_2” and “CO_2”. Next, students must make inferences about additional features, such as the effect of lighting on two different life forms that are placed in separate liquid-filled test tubes. Finally, students must make further inferences about yet further features, such as the effect of lighting, over time, on two life forms in the same liquid-filled test tube. This progression follows the idea that stages in a task should be ordered so that earlier stages train students how to think about later (potentially more difficult) stages.

Moreover, as students move through such a progression, they are able to engage in both *vertical* and *horizontal interaction* (processes that are often key to successful videogame performance). In vertical interaction, a student incrementally develops better skills, while in horizontal interaction, a student’s skills stay relatively the same but the student obtains, through exploration, a rough idea of what the various skills are and how they might be used (Goto 2003; Gee 2008). As Gee (2008) notes, “Horizontal experiences look like mucking around, but they are really ways of getting your feet wet getting used to the water and getting ready eventually to jump in and go swimming”. Both vertical and horizontal interaction play a role in ONPAR tasks. Students participate in vertical interaction when they progress through various stages of a task and see how the latter stages build upon aspects of the former stages. Students also have the option to participate in horizontal interaction; e.g., by freely moving forward and backward through the stages of a task, by positioning the cursor atop rollover icons, or by listening to spoken versions of written texts.

In order to facilitate both vertical and horizontal interactions, ONPAR tasks often encourage a process that is common in videogames: i.e., what Gee (2003) calls a cycle of “probing, hypothesizing, reprobing, and rethinking”. During the cycle, a player first probes the virtual world by interacting with and engaging the surrounding environment (e.g., the player searches a room, finds and opens a chest, and picks up a health kit). Then, on the basis of this probing, the player forms a hypothesis about the meaning of something in the environment (e.g., the player hypothesizes that finding a health kit means that there will be a difficult section on the other side of the door). The player then reprobes the environment with that hypothesis in mind (e.g., the player opens the door and passes into the next room expecting it to be a difficult section). Finally, the player uses the outcome of the reprobing as feedback to embrace or revise the hypothesis (e.g., the player finds that there is indeed a difficult section on the other side of the door and thus further embraces the initial hypothesis that finding a health kit has the meaning initially attributed to it) (for similar ideas, see Dewey 1933/1986; Schon 1987; Gee 1997).

A similar process occurs in many ONPAR tasks. A student probes the environment by pressing some button on the screen (e.g., the play button). Then, the student watches an animation and forms a hypothesis about which stimuli in the environment are most relevant to the task (e.g., the color of a liquid, the color of a rubber bulb on a dropper). Next, the student reprobes the environment (e.g., by pressing the replay button, or by moving the cursor atop a rollover icon or underlined word). Finally, the student uses the outcome of reprobing to embrace or revise his or her hypothesis (e.g., supposing that the student moves the cursor atop a rollover icon, he or she might notice relevant stimuli that are now highlighted on the screen, and can then use this as feedback to embrace or revise the initial hypothesis about which features in the animation are most relevant to the task).

On a more general level, ONPAR tasks, like good games, can be described as "well-defined problems nested within ill-defined problems" (Steinkuehler 2013). Well-defined problems have clear criteria and boundaries; they are constrained in such a way that it is clear whether one has provided a correct answer. Many problems that are relevant to gameplay, however, are not well-defined. In particular, there are also ill-defined problems; i.e., problems to which there are no determinately correct answers, but to which a player can nonetheless respond more or less skillfully (e.g., in the way that one can respond more or less skillfully to the issue of when to use a healing potion, or of whether to hide from a security camera or try to disable it).

ONPAR tasks can likewise be described in terms of well-defined problems nested within ill-defined problems. In a biology task, for instance, students are presented with two liquid-filled test tubes under light. One test tube contains a plant and the other contains an animal. Students must then infer, for each test tube, whether the liquid will turn blue (indicating O_2) or yellow (indicating CO_2). Along these lines, the task can be understood as a well-defined problem, e.g., in that students who indicate that the test tube with the plant will turn blue are correct, while those who do not are incorrect. However, the task can also be understood in terms of an ill-defined problem, e.g., in that students must decide with which presentation modes they want to engage. For instance, students can choose to focus solely on the animation without attending to the written or spoken instructions, or they can attend to the written and spoken instructions in conjunction with the animation. Similarly, they can replay a previous animation, replay the current animation, look ahead to a subsequent animation, or move the cursor atop a rollover icon or underlined word in order to highlight relevant stimuli or view a demonstration of how to perform a given action. The problem of determining which of these choices should be carried out is ill-defined, since there is no determinately correct answer to the problem. However, students can still aim to respond more or less skillfully to such problems, in part, by taking into account their abilities, prior knowledge, and surrounding context.

4. Conclusion

To summarize, ONPAR methods are designed to open up avenues by which students can better convey their content knowledge and skills. In this paper, we examined a large variety of studies from cognitive science to help explain how and why

ONPAR methods are likely to succeed at opening up such paths. To do this, we began by considering how ONPAR's use of assisted open-ended questions and instructions helps provide access to the ways in which various concepts, propositions, strategies, and procedures are related to one another in students' long-term memory. We then looked at how ONPAR uses multiple presentation modes to activate relevant knowledge schemas, and thereby to facilitate student performance, in those with relevant content knowledge and skills. Next, we focused on how the specific organization and design of ONPAR stimuli serve to facilitate efficient processing of relevant information; e.g., through the use of spatial and temporal proximity, the use of conversational versus formal language, the exclusion of interesting yet extraneous material, and the use of explicit signals to guide students' attention to relevant stimuli. Finally, we looked at various concepts from the videogame literature in order to further situate ONPAR among other dynamic multimodal environments.

This paper paves the way for a closer examination of several questions. For instance, how do the advantages of multiple presentation modes in learning and assessment relate to the disadvantages of redundant information (e.g., information that is presented in both written and spoken form)? Second, might the problems associated with redundant information disappear if students were not only presented with redundant information but were also given the option to *choose* the presentation modes through which they receive information? Third, if students can choose the presentation modes through which they receive information, is it more advantageous for them to choose just one or two presentation modes rather than most or all of the presentation modes? Fourth, do students inherently know which presentation modes are best suited to their learning and testing styles? Fifth, are the presentation modes with which a student prefers to work the same as those with which the student performs best? It is to such questions that we turn in future research.

References

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, 94: 416–427.

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology,* 30: 117–139.

Baddeley, A., & Della Sala, S. (1996). Working memory and executive control. *Philosophical Transactions of the Royal Society of London: Series B. Biological Sciences*, 351: 1397-1404.

Baddeley, A. D., & Logie, R. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah, eds., *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge, UK" Cambridge University Press, pp. 28-61.

Bajraktarevic, N., Hall, W., and Fullick, P. (2003). Incorporating learning styles in hypermedia environment: empirical evaluation. In P. de Bra, H. C. Davis, J. Kay, & M. Schraefel, eds., *Proceedings of the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems*, Nottingham, UK: Eindhoven University, pp. 41-52.

Baker, E. L. (2009). The influence of learning research on the design and use of assessment. In K. A. Ericsson (Ed.), *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 333-355). New York: Cambridge University Press.

Carr, T. G., & Kopriva, R. J. (2013). *Greater than the sum of its parts: Theoretical and conceptual underpinnings of the ONPAR methodology for conveying meaning through multiple, multi-semiotic representations*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

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

Clark, A. (1989). *Microcognition: Philosophy, cognitive science, and parallel distributed processing*. Cambridge, MA: MIT Press.

Clark, D. B., & Martinez-Garza, M. (2012). Prediction and explanation as design mechanics in conceptually integrated digital games to help players articulate the tacit understandings they build through game play. In C. Steinkuehler, K. Squire, and S. Barab, eds., *Games learning and society: Learning and meaning in the digital age*. New York: Cambridge University Press, pp. 279 – 305.

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

Cole, N. S., & Moss, P. M. (1993). Bias in test use. In R. L. Linn, ed., *Educational Measurement* (3rd ed.). Phoenix, AZ: Oryx, pp. 201-220.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agent in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, 94: 428–434.

Dehn, M. J. (2008). *Working memory and academic learning: Assessment and intervention*. Hoboken: Wiley.

Dewey, J. (1933/1986). *How we think. A restatement of the relation of reflective thinking to the educative process. John Dewey. The later works (Vol. 8).* J. A. Boydston. Boston: Heath.

Elman, J. (1991a). Distributed representations, simple recurrent networks and grammatical structure. *Machine Learning* 7: 195-225.

Elman, J. (1991b). *Incremental learning, or the importance of starting small*. Technical Report 9101, Center for Research in Language, University of California at San Diego.

Elman, J. (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, 48: 71-99.

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering Education*. 78: 674–681.

Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer ed., *Cambridge handbook of multimedia learning*. New York: Cambridge University Press, pp. 117-133.

Ford, N., & Chen, S. Y. (2001). Matching/mismatching revisited: an empirical study of learning and teaching styles. *British Journal of Educational Technology*. 32: 5–22.

Garner, R., Alexander, P., Gillingham, M., Kulikowich, J., & Brown, R. (1991). Interest and learning from text. *American Educational Research Journal*, 28: 643–659.

Gee, J. P. (1992). *The social mind: Language, ideology, and social practice*. New York: Bergin & Garvey.

Gee, J. P. (1997). Thinking, learning, and reading: The situated sociocultural mind. In

D. Kirshner and J.A. Whitson, eds., *Situated cognition: Social, semiotic, and psychological perspectives*. Norwood, N.J.: Lawrence Erlbaum, pp. 235-259.

Gee, J. P. (2001). Progressivism, critique, and socially situated minds. In C. Dudley-Marling and C. Edelsky, eds., *The fate of progressive language policies and practices*. Urbana, IL: National Council of Teachers of English, pp. 31-58.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave-McMillan.

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (2nd ed.). New York: Palgrave Macmillan.

Graf, S., & Kinshuck (2008). Technologies linking learning, cognition, and instruction. In J. M. Spector, M. D. Merrill, J. van Merriënboer, & M. P. Driscoll, eds., *Handbook of Research on Educational Communications and Technology* (3rd ed.). New York: Taylor and Francis, pp. 305-316.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, 3: 1–40.

Guida, A., Gobet, F., Tardieu, H., & Nicolas, S. (2012). How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: a two-stage framework. *Brain and Cognition*, 79: 221-44.

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89: 92–102.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, 90: 414–434.

Hays, T. A. (1996). Spatial abilities and the effects of computer animation on short-term and long-term comprehension. *Journal of Educational Computing Research*. 14: 139–155.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, 17, 329–433.

Jiang, Y. V., Swallow, K. M., & Rosenbaum, G. M. (2013). Guidance of spatial attention by incidental learning and endogenous cuing. *Journal of Experimental Psychology: Human Perception & Performance*, 39: 285-297.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of Individual Differences, Learning, and Instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

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

Kersten, A. W., & Earles, J. L. (2001). Less really is more for adults learning a miniature artificial language. *Journal of Memory and Language*, 44: 250-273.

Kirriemuir, J., & McFarlane, A. (2004). Literature review in games and learning. Retrieved June 06, 2014, from http://archive.futurelab.org.uk/resources/documents/lit_reviews/Games_Review.pdf.

Kress, G. (1985). *Linguistic processes in sociocultural practice*. Oxford: Oxford University Press.

Kress, G. (1986). *Before writing: Rethinking paths into literacy*. London: Routledge.

Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Edward Arnold.

Lohr, L. L., & Gall, J. E. (2008). Representation strategies. In J. M. Spector, M. D. Merrill, J. van Merriënboer, & M. P. Driscoll, eds., *Handbook of Research on Educational Communications and Technology* (3rd ed.). New York: Taylor and Francis, pp. 85-96.

Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. Mayer, ed., *Cambridge handbook of multimedia learning*. New York: Cambridge University Press, pp. 147–158.

Masson, M. E. J., Bub, D. N., & Lalonde, C. E. (2010). Video-game training and naïve reasoning about object motion. *Applied Cognitive Psychology*, 25: 166-173.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology*, 93: 377–389.

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

Mayer, R. E. (1997). Multimedia learning: are we asking the right questions? *Educational*

Psychologist. 32: 1–19.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.

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

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

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

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., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, 96, 389–395.

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

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

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

Mayer, R. E., Moreno, R., Boire M., & Vagge S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, 91: 638–643.

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

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology*, 95: 419-425.

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

McClarty, K., Orr, A., Frey, P., Dolan, R., Vassileva, V., & McVay, A. (2012). A literature review of gaming in education research report. *Pearson*. Retrieved June 06, 2014, from http://researchnetwork.pearson.com/wp-content/uploads/Lit_Review_of_Gaming_in_Education.pdf.

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.

Moreno, R. (2006). Does the modality principle hold for different media? A test of the method affects learning hypothesis. *Journal of Computer Assisted Lear*ning, 22, 149–158.

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

Moreno, R., & Mayer, R. E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology*, 92, 724–733.

Moreno, R., & Mayer, R. E. (2002a). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, 94, 598–610.

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

Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology*, 96, 165–173.

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review,* 19, 309-326.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177–213.

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.

O'Neil, H. F., Mayer, R. E., Herl, H., Thurman, R., & Olin, K. (2000). Instructional strategies for virtual environments. In H. F. O'Neil & D. H. Andrews, eds., *Aircraft training: Methods, technologies, and assessment*. Mahwah, NJ: Erlbaum, pp. 105–130.

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

Pellegrino, J. W., Chudowsky, N., & Glaser, R., eds., (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Schon, D. A. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass.

Squire, K. (2012). Designed cultures. In C. Steinkuehler, K. Squire, and S. Barab, eds., *Games learning and society: Learning and meaning in the digital age*, Cambridge: Cambridge University Press, pp. 10 – 31.

Steinkuehler, C. (2013). Lecture. *Video Games and Learning*. Lecture conducted from University of Wisconsin, Madison.

Stull, A., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology*, 99: 808–820.

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. Mayer, ed., *Cambridge handbook of multimedia learni*ng. New York: Cambridge University Press, pp. 159-167.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory (Explorations in the learning sciences, instructional systems and performance technologies)*. New York, NY: Springer.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General,* 119, 176-192,

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

Wey, P. and Waugh, M. L. (1993). The effects of different interface presentation modes and users' individual differences on users' hypertext information access performance. Paper presented at the *Annual Meeting of the American Educational Research Association*, Atlanta, GA.