Original article

doi: 10.1111/j.1365-2729.2008.00299.x

Designing multimedia learning environments using animated pedagogical agents: factors and issues

H.L. Woo

Learning Sciences and Technologies Academic Group, National Institute of Education, Nanyang Technological University, Singapore

Abstract Animated pedagogical agents (APAs) are known to possess great potential in supporting learning because of their ability to simulate a real classroom learning environment. But research in this area has produced mixed results. The reason for this remains puzzling. This paper is written with two purposes: (1) to examine some recent research and organize the findings in terms of classroom characteristics, and (2) to discuss and reveal any uncovered issues pertaining to the findings and provide input whenever possible. A framework formed by using APA characteristics, APA presentation, and learners' characteristics is used to analyse past research findings. The findings from the analysis reveal that because APAs are regarded as social members similar to humans, they are more effective in engaging learners in environments that require social communication and interactions. They therefore produce more definitive results in terms of affective gain and group learning. But such conditions also impose greater demand on designers to create more complex learning environments that can provide interactions with several agents and yet maintain a mode of communication that is pedagogically effective. The challenges for creating such environment include using an agent's gestures to duplicate its speech in instruction, which is usually uncommon in human practice, overcoming the needs to use input–output interface for communication and taking into consideration the possible influence of the learners' characteristics such as their sensory preference. This paper suggests that APAs' application in instruction should be seen in the light of affordances and be designed within its own practical limits.

Keywords Agent, affordance, cognitive overload, modality, multiplicity, persona effect, sensory preference.

Introduction

The use of multimedia has revolutionized the way people learn, especially for online learning. It uses the idea of utilizing as many human senses as possible to reach full learning engagement. The five senses are seeing, hearing, feeling, tasting, and smelling. Practically, only seeing, hearing, and feeling are used in most online multimedia learning because the experience of taste and smell are not easily achievable using online technology. Research has found that by exposing learners to only hearing, about 5% of information presented may be retained, but when hearing and seeing are combined in a learning process, such as showing a picture with voice explanation, learners can retain up to 20% of the presented information. More so, if demonstration is included in the presentation, then retention can increase

Accepted: 10 September 2008
Correspondence: Woo Huay Lit, Learning Sciences and Technologies Academic Group, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616. Email: huaylit.woo@nie.edu.sg

to up to 30% (Magennis & Farrell 2007). Because retention is a vital component in a learning process, this finding suggests that multimedia that utilizes both sight and sound heavily in its instruction is a good way to help learners learn.

Despite the cited advantages of using multimedia in online learning, there remains a setback that the online community is trying to overcome: the lack of human factors during learning online with the multimedia. Although multimedia engages many forms of media, such as a video, which contains images and voice of human beings, the human presence adds merely to the content domain of learning and provides little human intervention to aid the pedagogy. Hence, the function of multimedia is still very much confined to content enhancement.

More than a decade ago, online learning had started to engage new technology to alleviate the non-human-touch problem by using, for example, videoconferencing and 'electronic whiteboard' in online courses (electronic whiteboard is an electronic sketchpad where online users can share what they write or draw during an online conferencing). These devices can help improve human-to-human communication, but they cannot be used anytime without prior arrangement for the group to meet at a specific time. In other words, just-in-time guidance is limited and so is human presence. Therefore, the search for a better means of human presence in online instruction will have to continue.

Of late, a new branch of multimedia using lifelike characters to aid teaching and learning has emerged. These characters are able to mimic human characteristics, including voice, gesture, and body movement. They are called animated pedagogical agents (APAs). They usually coexist with the learning materials on a computer screen so that they can be regarded as an integral member of the learning environment. As such, they are normally designed to take up a central role, such as tutoring or facilitating.

Since the introduction of APAs in online learning, a growing body of research has tried to ascertain the impact of APAs on learning gains. Many were interested in how APAs affect the affective and cognitive domains of learning. A good number of the studies that attempted to establish a relationship between APA learning and motivation found a positive effect and attributed it to the *persona* effect brought about by the lifelike behaviour of the agents (e.g. Lester *et al*. 1997; Mulken *et al*. 1998; Atkinson 2002; Moundridou & Virvou 2002; Baylor & Ryu 2003). The *persona* effect is a result of anthropomorphizing the agent, which gives the learner a feeling of being accompanied by a real person during the learning process. This is good testimony to learning enhancement in the affective domain of learning and an obvious solution to the lack-of-human-touch problem. But results for learning gains in the cognitive domain have not been equally overwhelming. Research in this area has produced mixed results and the findings are inconclusive. The question is, why does APA, as a supporting means to human teaching, fail to produce positive results to achievement gains? Could it be that the role of an APA now resembles that of a human so it is facing the same kind of complexity as a human tutor? Or perhaps the presence of an APA has added more dynamism on a computer screen, which has complicated the way we perceive information?

There is a need to look closer at these issues. This paper is written with the aim of providing some input to the discussion. It looks into past research and tries to explain the phenomena by using common classroom characteristics. It hopes to bring to light a more holistic perspective that will aid understanding of the problems and factors affecting APA learning environments.

Categorizing APA research

The uniqueness of APA research is that it deals with virtual objects (agents) that have a tendency to approach reality. The close resemblance between an APA and a human has made the rendering of APAs in instruction very complex and difficult. For this reason, findings from APA research need a structure to organize and analyse. The following argues that using an APA to deliver instruction is similar to a teacher teaching in a classroom but an APA, unlike a real teacher, cannot exist alone, so it needs to be situated in a learning system that provides support for its instructional rendering. Therefore, similarity to a classroom and the system's architecture will be used as two major considerations in forming a framework for categorizing and organizing APA research to be used in this paper.

Similarities to a classroom situation

APAs are commonly used in situations that resemble those of a real teacher in a classroom, except that an

APA learning environment is technologically driven and frequently carried out via online delivery. This is because an APA-based learning system is considered an extension of the more traditional system called the intelligent tutoring system (ITS) (Chou *et al*. 2003; Dos Santos *et al*. 2002; Gulz & Haake 2006; Wissick 2002). An ITS is designed to provide an individualized learning condition with the computer acting as a personalized tutor. With the use of a humanlike APA, the tutoring role is then assigned to the APA instead of a machine, thereby making the learning experience more natural and realistic. Because the APA has to assume the role of a human tutor, it is expected to possess most, if not all, of the characteristics that are required by a good human tutor. So tutoring by an APA is very similar to teaching by a real teacher except that for the former, learning takes place normally in a fixed environment that encompasses only the computer and the learner whereas for the latter, learning takes place in a classroom that has more learners and changeable conditions. Despite this environmental difference, the requirements for a good tutor or teacher in both cases remain the same. Seen in this light, the role of an APA implies that the analysis of APA research would be best carried out by matching the APA attributes with those for a real classroom.

Based on the APA's architecture from a design perspective

An APA is an artefact; to fabricate a humanlike APA will require a system capable of handling the APA's behaviour and the interface dealing with the input and output between the APA and the user. Normally, such a system can be treated as being made up of two components: the front-end and the back-end. The APA's behaviour including the interaction with the users is commonly managed by using artificial intelligence (AI) (Nunes *et al*. 2002) which resides at the back-end of the system. AI has the advantage of managing 'reasoning' that takes control of all behavioural function of the system (Ruttkay *et al*. 2004). The front-end of the system usually consists of the APA's display and its embodiment with other media that form the presentation for the learning. With this view, the study of APAs for learning may likewise be divided into two areas of focus: one focuses on how the APA should appear and behave in order to provide optimal learning, and the other concentrates on how to materialize the optimization through the technical means. The former deals mainly with the pedagogy and the latter handles the technicality to support the pedagogy. Separating functionally the instructional aspect of design from the technological provisions allows a clearer discussion on the factors and issues affecting APA-based learning environments without the complication of technology.

Because the APA presentation at the front-end involves the design of embodying an APA in instruction and the process usually requires consideration of integrating visual and auditory modalities, the discussion for this aspect of design will entail knowledge of multimedia principles. Multimedia is a field that deals with the presentation of material using both words and pictures where words are in written or in spoken forms and pictures include both static and dynamic graphics, photos, animations, and videos (Mayer 2001). Therefore, multimedia will be used as a platform for designing APA learning environments in this paper.

A framework to guide the research review

In a real classroom, there are four most fundamental elements that can affect learning – the teacher, the teaching, the students, and the physical environment (Hetherington & Parke 1993). For an APA to emulate a good teacher, it has to have good APA characteristics (similar to good teacher characteristics) and a good way of presentation (similar to good teaching). It must know that its instruction is also affected by the students' characteristics and the physical environment in which the lesson is taking place. With the exception that the APA physical environment is quite invariant and confined to working with a computer and its peripherals, the other three elements provide a good basis for forming a framework that helps to organize research findings according to some criteria. In this respect, a framework called APL is formed based on the classroom attributes. The meaning of the acronym is given below:

A represents 'APA characteristics',
P represents 'Presentation by the APA', and
L represents 'Learners' characteristics'.

The APA framework represents the three key areas of focus for the discussion of this paper. In the following sections, research findings pertaining to APA-based learning will be reviewed and discussed according to the three areas from the framework. However, because

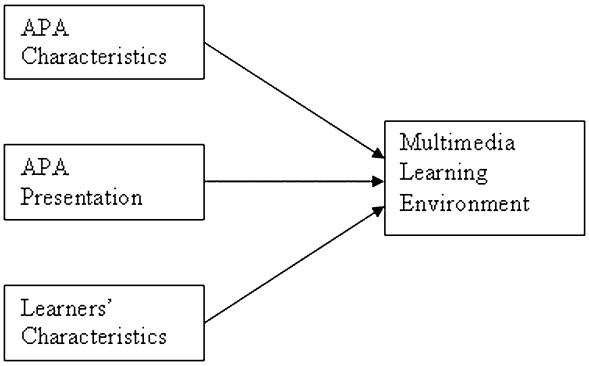

Fig 1 The APL framework where *A* is animated pedagogical agent (APA) characteristics; *P* is Presentation by the APA; *L* is Learners' characteristics.

of the vast number of different issues that were generated in the research, only controversial issues and factors are highlighted. (Fig 1)

APA Characteristics

The central issue on APA characteristics is whether an APA should exist like a human or remain as an artefact. Closely related to this issue are two most frequently unanswered areas – anthropomorphism and the appearance of an APA.

Anthropomorphism

Many studies have focused on what APAs can provide in relation to enhancing the learning experience. Examples are Dehn and van Mulken (2000); Johnson *et al*. (2000); Baylor (2002); Clarebout *et al*. (2002); Moundridou and Virvou (2002); White *et al*. (2002); Baylor and Ryu (2003); Buisine *et al*. (2004); Moreno (2005); and Woo and Wang (2005). These studies found that by rendering agents with lifelike features, such as facial expression, deictic gestures, and body movement, they are able to emulate any real-life character and give rise to what is called a *persona* effect. A persona effect is a result of anthropomorphism derived from believing that the agent is real and authentic (Mulken *et al*. 1998; Baylor & Ebbers 2003).

The humanizing of the agents, according to Baylor and Ryu (2003), is able to provide the emotional connection between the agents and the learners, thus increasing the learner's enjoyment of the learning process. This supports findings from cognitive science that learning is most effective when done under a condition when fear or anxiety is low (Ormrod 1995, pp. 425–430; OECD 2002, pp. 24–26). Having a lifelike companion brings familiarity to the learner because the learning condition resembles that of a teacher and a student in a face-to-face classroom. As a result, the fear and anxiety levels can be kept to the minimum. This is a significant contribution to online learning because a common criticism about learning using a computer is that the learning is 'cold' and 'aloof' (Chang & Fisher 2003). Thus, APAs help bring back the human touch and make online learning a less threatening experience.

On whether anthropomorphism can also promote learning, there are fewer conclusive results (e.g. see Mulken *et al*. 1998; Moundridou & Virvou 2002; Prendinger *et al*. 2003). The main reason is that anthropomorphizing an agent is a complex process and measuring the degree of learning can also be just as complicated. The effect of anthropomorphism cannot be viewed simply as presence or absence of an agent but must be looked upon as an integration of the agent's behaviour with the instructional approach; it is therefore an amalgam of pedagogy and interface design. For this reason, it is difficult to study in isolation the effect of each contributing factor under an empirical condition (Dehn & van Mulken 2000). This contrasts greatly with the result on learning experience offered by APAs because experience by itself is a subjective construct and can be measured by one's overall feeling. This might explain why there are more conclusive results on the affective aspect of APAs than the cognitive benefits obtainable from APAs.

Despite the abovementioned difficulties for designing APA experiments involving cognitive measurement, some studies have taken steps to simplify the anthropomorphism process and use a tighter control on the empirical setting. One of them is from Beun *et al*. (2003). Beun *et al*. carried out an empirical study on whether anthropomorphism using conversational agents can affect memory performance. They created three conditions, namely, realistic, cartoon, and absence. The realistic condition has a human face that communicates and reads a short story to the participants; the cartoon condition has a full body of a gorilla and communicates and reads the same story in the same manner; and the absence condition has no agent at all and uses only words for communication and showing the same story. All three conditions have similar settings

except for the agent's appearance, which is used as a variable. It should be noted that all three conditions use a word balloon to display the verbalization.

The result of the study showed that both realistic and cartoon conditions can produce a positive effect on memory performance, but the result is negated by the questionnaire responses from the participants in that they considered anthropomorphism is due only to the realistic agent but not to the cartoon-like agent. The study therefore posits that conversational agents have the potential to produce a positive effect on memory performance but the effect may not be entirely due to anthropomorphism. Although the result helps to establish a causality effect, it fails to provide any explanation for it. In addition, the study is likely to suffer a pedagogic design problem. This is because the presentation uses spoken words that are redundantly repeated as text in the word balloons. This is a violation of the redundancy principle (Mayer 2001; Sweller 2002; Moreno 2005), and it also causes the split-attention effect to increase (Sweller 2002). The redundancy effect has a tendency of directing the learners' attention away from the narration and the split-attention effect tends to distract the learners' vision from focusing on the agent. Because of this inherent problem, it is not sure whether the memory performance is actually due to remembering the text in the word balloon or to the narration from the agents. It is also not sure whether the learners have paid sufficient attention to the agents before they form their perception. Nonetheless, the study has helped to illustrate the difficulty in balancing the pedagogical requirements and the research demands when designing an empirical study involving pedagogical agents.

Appearance

The issue of appearance is less crucial for a teacher in a real classroom. This is likely that students in a real classroom have little choice over what their teacher should look like. But for an APA, because it is artificial, designers have a choice in the appearance of the agent to meet the learning requirements. Given this choice, some researchers, such as Walker *et al*. (1994); Beun *et al*. (2003); Buisine *et al*. (2004); Gulz and Haake (2006); and Buisine and Martin (2007), are interested to know whether an APA should be portrayed with a real human look or simply with an iconic feature such as a cartoon face.

Different researchers associate the appearance of an agent with different instructional effects. Briefly, the association may be linked to the learners' perception and the context in which the agent is situated.

Learners' perception of the agent

Welch *et al*. (1996) argue that agents with a pictorial image that has close-to-human look increase the level of involvement and sense of presence in a learning environment. This is echoed by Nass *et al*. (2000), who propose that the appearance of agents should be made to resemble that of the learners. However, McCloud (1993) holds a different view. McCloud argues that individuals see themselves in their own minds as iconic images but see others in a more detailed form, that is, as realistic images. Therefore, learning activities that involve identity recognition and social affinity will favour the use of iconic agents. McCloud also suggests that iconic characters have the advantage of being portrayed as having a less serious look but still exhibits characteristics in a realistic way.

Gulz and Haake (2006) extend their view to the role of an APA on the ground of McCloud's idea. They argue that when an agent plays a role of a teacher, because the role is external to the learners, it is seen as 'the other person' and therefore is better to be represented in a human form. But if the agent is to play the role of a learning companion, then it has the same status as the learners, and so is being seen as 'self', and is better represented in iconic form. However, such view is opposed by Buisine and Martin (2007), who cited Kohar and Ginn's recommendation (1997) that dramatized characters can display more exaggerated emotions than realistic humanlike agents. They illustrated this effect by using a realistic agent but with a cartoonish broad smile in one of their earlier experiments (Buisine *et al*. 2004). The result was that this hybrid appearance (realistic image with cartoonish smile) tended to produce greater likeability. Another supporting finding is from Ruttkay *et al*. (2004), who argue that humanlike agents may be seen as more intelligent but non-humanlike ones can be more appealing and entertaining. Other related commercial uses of iconic characters include the Disney and the Sesame Street programs. The characters used in these programs are iconized, and even with exaggerated expressions, they are long known for their ability to engage audience, both young and old.

There are also research findings that found a neutral effect on the 'iconic vs realistic image' issue. For example, Moreno *et al*. (2001) in a very detailed study compared two learning conditions: one condition had an agent with only an iconic feature of a person but without facial expression and gaze, and the other condition had a similar agent but with a full expressive video image. The study found that the two conditions do not produce any significant difference in terms of retention, transfer of learning, and participants' interest of learning. The results indicate that whether the agent image is iconic or realistic and whether the appearance is expressive or not do not affect their cognitive and affective outcomes. Moreno *et al*. explain that what actually influences the learning outcome is whether the agent is able to provide an interactive environment through communication. In fact Beun *et al*.'s (2003) study (as discussed earlier) also produces the same result despite a possible flaw in their pedagogical design.

From the preceding discussion, it seems that look alone is not perceived as an important element for instructional effectiveness, but rather it is the roles the agent plays and the types of learning tasks that give rise to the effect.

The context in which the agent is situated

Reeves and Nass (1996) in their 'media equation theory' suggests that humans have a tendency to ascribe social norms to whatever medium they are interacting with because humans regard the medium as a social actor. Being a social actor, the agent must perform social activities such as talking and reacting to learners' responses to provide interactions. It is very likely that it is the level of interaction, and not just the face alone, that determines the effectiveness of an agent to support learning (Badler & Allbeck 2001; Wang *et al*. 2008). This is analogous to a teacher in a real classroom situation in that most students judge a teacher's performance by his or her holistic teaching effort and not just by his or her looks. However, there is still a limit to the level of interaction an agent can provide. This is so because the interface by which an APA interacts with a learner limits the realistic effect experienced by the learners. Interacting with an agent in a computer is different from a real classroom where communication is spontaneous and bidirectional; and the modality can be either verbal or non-verbal or both. The present system of APA communication is mostly unidirectional and it favours the tutor-to-student style of communication (Baylor *et al*. 2005). When the communication is in the student-to-tutor direction, additional interface must be used. It needs a microphone to listen to the student's voice. In some cases when the APA is given more intelligence to read a student's body cue in conjunction with the verbalization, then a video camera must be used as well. The technology requirements just described put a limit to which an APA can function like a human. To make APA learning a reality, users must adapt to new changes. This is a challenging endeavour and will take time to achieve. Therefore, it seems that the research effort attempting to make APAs as real as possible may not prove to be fruitful.

APA presentation

Presenting an APA to facilitate learning, like teaching in a classroom, is the most crucial and difficult procedure to execute in the entire instructional process. Not only that instruction itself is a very trying business, it entails a repertoire of different events, such as showing resources, explaining content, and communicating with students to make it work. All these must be well synchronized and skilfully delivered to obtain desirable results. There is no best formula that can be used for achieving the desired instruction. Each learning scenario is unique by itself and requires a different instructional approach. In the same manner, each APA research study adopts a different instructional approach and uses different participants. Although attempts may be made to standardize the control for variables, such as establishing a common framework for research design (see Dehn & van Mulken 2000; Clark & Choi 2005), the instructional approach remains varied because of the different contexts for research and environmental settings. In other words, replication of research studies is rarely found and generalization of research findings remains difficult.

One of the ways to resolve this problem might be to analyse the instruction approach by parts. This means that the instruction may be analysed by seeing how an APA communicates and how it directs instruction. The reason why these two modes of instruction are chosen is that they are the most fundamental strategies to be found in any instruction. Issues concerning these two modes of instruction are presented in the following sections.

How APAs communicate the instruction?

From the standpoint of multimedia learning environment, an APA can communicate its instruction by two modalities – visual or auditory (Mayer 2001; Moreno 2005). The term modality means 'the sensory format in which information is displayed' (Clark 2005, p. 613). Modality has attracted the most concerns in designing multimedia presentations. One of the prominent theories that was highly quoted is the cognitive theory of multimedia learning (CTML) (Mayer 2001; Mayer 2005a,b), and most recently, the cognitive theory of multimedia learning with animated pedagogical agents (CTMLAPA) (Moreno 2005). Because the term 'multimedia' implies that information can be presented in a myriad of formats, such as text, graphics, animations, and videos, there is a natural concern for whether the different forms of display will compete for our limited sensory capacity.

Cognitive theories have informed us that our mind is only able to form a rough impression called *perception* from what we receive from our sensors (Ormrod 1995, pp. 194–195). Perception produces an initial representation quite different from what is actually displayed on the computer screen. This is because we have very limited sensory memory available to our eyes and ears to capture the entire information. So the eyes and ears select only a portion of the information deemed relevant and send it to our working memory (see Fig 2). As a result, perceived information usually results in incomplete representation. The extent to which the incomplete representation is formed depends on many factors, one of which is the dissimilar processing mechanisms that exist in the visual and auditory channels.

The processing mechanism of the two channels can be explained by the CTMLAPA model in Fig 2. The model shows that the route by which information enters our mind depends on whether the information is in a visual format or in an auditory format. Visual information proceeds through the lower channel and auditory information proceeds through the upper channel. When a presentation consists of both visual and auditory items, such as a narrated animation, the voice portion is filtered by the ear sensory memory to form a *propositional representation* (Schnotz 2005) called the verbal mental model in the working memory. Similarly, the animation portion is filtered by the eyes to form a visual perception called the pictorial mental model in the working memory. The two mental models do not exist in a disjunctive manner. In fact, they share their attributes. For example, the pictorial model may comprise components of spoken words obtained from the narration; likewise, the verbal model may contain tinges of images gathered from the animation. The two channels produce a temporary 'hybrid' model until it is integrated with the learner's existing schema (Schunk 2000) in the long-term memory to form a more refined representation called 'knowledge'.

It can be seen that, although the cognitive process is a complex one, the entire process depends quite significantly on the early stage of perception. Therefore, the way the visual and auditory materials are displayed would have a crucial impact on the perception process.

For an APA presentation, there are at least three forms of visual elements that our eyes have to handle – the written words, the agent image (including its gesture and facial expression), and other pictures that reside in the background. Therefore, this has the potential to create a heavy demand on the visual channel (Beun *et al*. 2003). When the channel fails to cope with this demand, it creates what cognitive scientists called the *cognitive overload* effect (Mayer & Moreno 2003; Low & Sweller 2005, p. 148). When this happens, our learning efficiency is compromised.

However, not all APA-based instruction has this inherent cognitive overload problem. In fact, Sweller

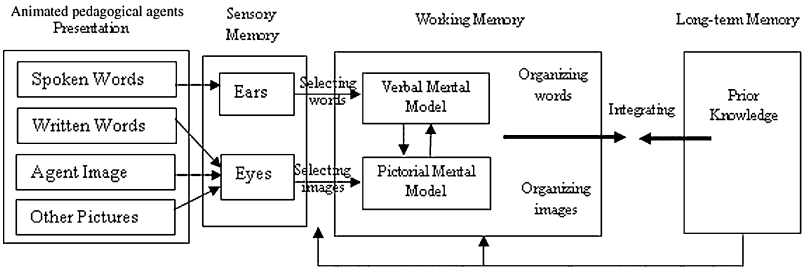

Fig 2 Cognitive theory of multimedia learning incorporating animated pedagogical agents (APA). Adapted from Moreno, 2005, p. 511.

(2002) pointed out that only instruction that has high element interactivity is likely to suffer the cognitive overload effect. Element interactivity means materials that contain learning components that are interrelated to one another and cannot be learnt independently. Sweller's idea is that when a learner learns with low element interactivity content, he or she probably has sufficient cognitive capacity to cope with the demand, and hence even when the instructional design is not conducive, its effect is not readily felt. But when a learner learns with high element interactivity content, his or her cognitive capacity is likely to be overtaxed, and so a poorly designed instruction will probably aggravate the situation, causing a gross cognitive overload to be experienced.

To use instruction to reduce cognitive overload, APA designers frequently turn to the two channels residing in the cognitive mind and try to optimize their usage. Because there is a higher demand in the visual channel for APA-based instruction, designers try to transform some of the visual presentation into auditory form so that learning can take place optimally in both channels. As an example, instead of using text to explain a certain procedure, make the agent verbalize the explanation and use minimum gestures. Many research findings have endorsed the positive effect of this strategy (e.g. see Moreno and Mayer 2000; Atkinson 2002; Buisine *et al*. 2004). However, enhancing voice usage alone may not be always desirable. Research on human–computer interface has cautioned us about the limit to which verbal voice can be used. It pointed out that as speech rendered by computers has become more and more authentic (closer to real human voice), people's belief in the humanistic behaviour of the computer grows. As a result, learners begin to see the voice as semantically commanding, which could lead to negative feelings and unusual behaviour towards the speaker (Hapeshi 1993).

When speech is used in computer-based learning, research recommends that it should be considerably short to carry only the essential information. Long passages should be delivered in chunks of short sentences. However, because speech presentation is serialized, even with short sentences, the speech is somewhat 'transient distributed' over time (Hapeshi 1993, p. 181). The longer attention span adds extra burden to the auditory memory just as the APA's appearance adds additional demand to the visual channel. But using auditory material has one advantage over the visual counterpart, that is, auditory information tends to stay longer in our memory (Wickens 1992). With this, it seems that each channel has its own limitation and proper usage. Merely using auditory channels to complement visual channels may not always work out well. Perhaps, APA-based learning environments might be more suitable when modality preference is not a significant consideration. The issue of modality preference will be discussed in greater depth in the section 'Learners' Characteristics'.

How APAs direct the instruction?

APAs can offer more advantages than humans for instruction in two ways, that is, by using multiple agents to create a social context for learning and by fabricating speech combined with gestures in proper coordination to optimize presentation. Real humans often have problems coordinating in a social setting and also require good skills and experience to apply the right use of speech and gestures during instruction. But because APAs are fully programmable objects, they are virtually free from these problems.

The use of agent multiplicity for social learning

Believers of social context learning tend to treat APAs as natural candidates for construction of a socially mediated learning environment. They see a connection between agent multiplicity and sociality in a learning community (Hietala & Niemirepo 1998). This is because using humanlike agents that have the ability to display social behaviour, especially in a group setting, has a greater tendency to communicate with learners more naturally (Beun *et al*. 2003). Some researchers regard agent multiplicity as a phenomenon whereby agents are situated in different machines to collaboratively learn and interact together with a learner; while others see it as a situation in which a few agents coexist in the same machine to share a learning experience with a human learner (e.g. Chou *et al*. 2003; Gulz & Haake 2006); yet some view agent multiplicity as an environment in which multiple agents play multiple roles (e.g. Baylor 2002; Baylor & Chang 2002; Baylor & Ebbers 2003). Research in this field found that two agents playing two different roles are better than one. This is because when different roles are performed by a single agent, the roles are not always in good match with the

agent's character. Unless the roles are so well portrayed, it usually leads to confusion and distraction.

Despite the many benefits obtained from using multiple agents, there remain issues such as how the agents are going to coordinate with each other to collaboratively present the content. The scenario is similar to team teaching in a classroom. Agent-mediated team teaching has one distinct advantage over human team teaching in that agents do not have innate psychological issues such as incompatible personality among team members. This is because agents are programmable. But to program agents as social characters, the programmer needs to have sufficient social engineering concepts to make the social context pedagogically viable. A good illustration of this is the 'eShow Room' project done by Krenn *et al*. (2002). In this project, they used lifelike agents to emulate a car sales scenario through a narrative context similar to that of TV commercials. A potential buyer obtains information about the car that he or she is interested in by watching and listening to the scene played by two agents; one acts as a salesman and the other as a buyer. The human buyer learns from the interplay between the agents and makes an informed choice about whether to buy the car or not. The whole learning process is set in a social environment using familiar day-to-day phenomena as the context for learning. Such a presentation involves the concept of vicarious learning and some knowledge of dramatization. Vicarious learning has its root in social cognitive theory, which emphasizes the tripartite interaction between behaviour, environment, and the individual factors in a learning process (Eggen & Kauchak 2001, pp. 238–239). The individual factors, which is the buyer's intention to buy, coupled with the enticing environment enacted by the agents, subtly produces a positive perception in the buyer's mind, which could, according to CTMLAPA, combine with the buyer's prior impression about the car, resulting in a favourable decision.

The eShow Room depicts another crucial but yet difficult-to-solve problem in the presentation of APA. APA presentation does not involve only simple pedagogy; it also needs peripheral knowledge of other disciplines, such as the aforesaid dramatization to create the sales scene, in particular, in situations where several APAs are simultaneously involved. The synergy of such effort is demanding, and unless it is done well, the impact rendered by APA learning will always remain questionable.

The use of speech–gesture cooperation for presentation

Learning is a social process that is mediated by communication normally in the form of speech and body movements (Gulz & Haake 2006). In fact, the bulk of the APA's instruction lies in its ability to explain and provide guidance through its speech and gestures. But not all speech and gestures have pedagogical values. For example, the gesture one uses in a conversation may be a result of body mannerism and not necessarily for conveying any instructional message. Therefore, what should be the right gesture that would go well with a right speech becomes a crucial consideration when designing an agent presentation. Studies attempting to find out what impact agents' speech and/or gestures can have on learning include, for example, Craig *et al*. (2002); Baylor *et al*. (2005); Rickel and Johnson, (2000); Lester *et al*. (2000); Piwek (2003); and Lusk and Atkinson (2007). These studies provide good discussion on issues related to an APA's conversational behaviour, but few of them actually try to determine empirically if there is a right combination for gesture to coordinate with speech to achieve an optimal communicative effect.

Buisine and Martin (2007) have conducted a useful study on the effects of speech–gesture cooperation in animated agents. They compared two modes of speech–gesture cooperation: redundancy and complementarity. Redundancy refers to duplication of information in speech and gesture, for example, pointing to a triangle and saying out in words 'the triangle'. Complementarity means the information is split into two with one delivered by speech and the other by gesture; both speech and gesture must be delivered in a way that they complement each other. Buisine and Martin's study found that redundancy is better than complementarity in terms of improving recall in verbal information, quality of presentation, and social perception of the agents (cf. Buisine *et al*. 2004). This is an interesting finding in that such behaviour is not commonly found in human communication because human communication is not restricted to only one mode of expressions but to both (redundancy and complementarity). Nonetheless, the finding appears to give support to the fact that agent-directed instruction may be more advantageous than human-directed instruction because designers have full control of the agent's behaviour (speech and gestures) and to work the agents according to some desirable pedagogies.

Note that Buisine and Martin's findings are obtained from the perspective of a designer. But the same may not apply when the results are viewed from a learner's perspective. This is so because there is still the subtle unintentional message that comes along with every agent's body movement, called body language, that conveys some unnoticed information. The unnoticed body language may be perceived as negative by some learners on some occasions. Pease (1987), quoting Albert Mehrabian's findings, noted that when one receives a message, about 7% of it is received verbally (words only), 38% of it is received vocally (including tone of voice, inflection, and other sounds), and 55% of it is received non-verbally (p. 9). These non-verbal messages include gestures, postures, position, and distances. Thus, not all of these body manifestations can be enacted by an agent in a presentation, and not all are properly investigated by research. Take for example the tutoring system 'Steve'(Rickel & Johnson 2000). Steve is an intelligent tutor meant to teach procedural knowledge about ship machines. It is designed to display only an upper-half body so that it does not cause too much obstruction to its background display. Although Steve is known to be effective in its teaching, it is not known if the partial display of its body has impaired its ability to convey full body language such as postures. Reeves and Nass (1996) pointed out that humans like to ascribe social rules to the media they are associated with. They also mentioned that the closer the media exhibit similarities to human social structure, the stronger the social rules would apply to the media. If an agent is to behave like a human, the more human social expectations will be imposed on the agent. Therefore, it seems that an agent's behaviour is correlated to the user's social expectations. The more the agent wants to approach real human behaviour, the more the user would expect the agent to be spontaneously 'smart' to respond to social norms. If the agent is unable to match the expectations, the user will become less tolerant. Therefore, the effect appears to induce a vicious cycle in that the more the agent is approaching reality, the greater the expectations the users impose on the agent, and the effect reinforces itself repeatedly. This probably explains why iconic agents are preferred to realistic looking agents reported in some studies (e.g. Buisine & Martin 2007; McCloud, 1993) because iconic agents are subjected to fewer social norms.

Learners' characteristics

The learner's characteristic is an area that receives the least research attention. Probably, many people regard the design of APA learning as a technically oriented discipline and hence channel most of the energy to producing a good agent artefact, but forget about the users of the artefact.

The previous discussion on modality pointed out that an APA animated presentation has a tendency to impose more cognitive load on the visual channel than on the auditory channel because of its unique characteristics and the channels' limited memory capacity. But research did not discern whether the channels are 'trainable' to become better in handling one modality than the other. Research has found that some people develop better visual acuity than listening ability and vice versa. People tend to use their stronger sensor to process information in their daily life. For example, visual people prefer to jot down a newly given telephone number, whereas auditory people tend to repeat the number aurally to reinforce the retention (Smith 2007). Learning style theorists call this phenomenon the sensory preference for learning (Dunn & Griggs 1998).

The existence of sensory preference could imply a possibility that one's sensory channel may be developed to become more proficient than what it usually is. Low and Sweller (2005) cited an experiment about shadowing a typist on different tasks to demonstrate the effect of modality on learning. In this experiment, the typist was to perform two tasks – one was to do typing from a given text passage while simultaneously listening to an oral message, while the second task was to carry out a similar typing but with the text and oral passage in reversed order. The experiment found that the typist could shadow the first task properly but had difficulty performing the second. This experiment confirmed that attention for a task is related to familiarization and experience with that task (Ormrod 1995). For most typists, they are trained conventionally to type from a written prose. So a skilled typist does not need much attention when typing in this way. But when the typist is asked to perform the job under an unfamiliar condition, such as typing from an oral prose, a great deal of attention will be needed. As a result, the typist has little 'spare' attention for any concurrent task and is unable to do the second task proficiently.

The above example points to the fact that training produces familiarity and hence determines the amount of attention needed for a task. Most training produces proficiency in only one modality. So a person's modality preference will depend on his or her training. Most of us are brought up in an environment that uses our sight more than our hearing (Wong & Cheung 2003). According to a report by the Learning and Skills Research Centre, (LSRC 2004), about 60% of us are visual learners. This could imply that many of us are more proficient in processing visual information and are less comfortable with working with auditory materials. If this holds true, then many of us may not feel burdened by the visually demanding APAs because of our increased visual capacity. So it appears that APAs could benefit visual learners more than auditory learners. But when the APA begins to include voice into the presentation, the situation becomes more complex. This is because voice may not favour visual learners. This argument apparently contradicts some of the research findings that showed voice narration with animation is beneficial to learning (cf. Moreno & Mayer 2000; Atkinson 2002; Buisine *et al*. 2004). The discrepancy between the two observations may be resolved by considering the different extent by which people use their sensory preference. Rose and Nicholl (1997, pp. 90–96) in a study noted that about 70% of learners learn well in any of the given modality; the other 20% rely on only one modality to learn, and the remaining have learning difficulty irrespective of the form of modality. So it seems that modality consideration in APA learning is still important to many of us.

The above discussion reveals that effective APA learning cannot rely on good presentation principles alone; it must also ensure that the presentational modality matches coherently with the learner's characteristics. This is like a classroom teacher trying to adopt a style of instruction that appeals to the students' background. Hence, it is proposed that sensory preference be considered an important student characteristic that should be taken into account, together with other instructional factors, when designing an APA-based lesson.

Summary and discussion

The similarity between a human instructed learning environment and APA mediated learning environment has allowed a common platform to be used for the discussion. This involves seeing the APA learning in the light of its characteristics, presentation, and relationship with learners' characteristics. With reference to research findings used in this paper, the following issues have been identified and discussed:

On APA characteristics

Generally, there is a consensus among researchers that learning experience can be made more pleasant by the help of APAs. However, there is little evidence to show that performance is related to the anthropomorphic effect of APAs. Similarly, research cannot authenticate the positive benefit of using realistic looking agents for instruction as compared to its iconic counterparts but suggests examining other factors such as ability to communicate with learners and types of role an agent plays which may be more tangible in contributing a difference. Closely related to an agent's appearance is the context in which the agent is situated. This paper finds that because an agent-based learning environment needs an interface such as a keyboard for text input, a microphone for voice input, or even some wearable gadgets for tactile input, the interaction is always restrictive and unnatural. Therefore efforts to make an agent behave like a real human will always be limited by the need of another communicative device. This means that there is always a limit to which an agent can approach human reality.

On APA presentation

An APA, together with its gestures, has a tendency to overload a learner's visual capacity especially when the content to learn is already demanding. Although a common practice to alleviate this problem is to convert and present some of the visual information in auditory format such as voice, the use of voice has its own limitation because voice transmission takes a longer time to be perceptually understood so it could also demand excess attention from a learner.

The use of multiple agents to forge a socially mediated learning environment has shown to be operating in alignment with the nature of actual classroom learning where knowledge sharing is favourably the norm. Although agents are artefacts, which can be programmed to behave in a group setting, the learning may

not carry any pedagogical value unless the intricacy of group behaviour and the art of social learning can be properly infused and meaningfully embedded in the learning process.

One of the important concerns for an APA presentation is how to achieve a best speech and gesture cooperation. Far from common belief and usual human behaviour, the redundancy mode of speech–gesture cooperation seems to be able to produce better cognitive gain than the complementarity alternative. But one must be cautioned about the redundancy behaviour because it tends to miss out in the process a very important and yet not easily reproducible action – the human body language. The body language helps to convey hidden meanings, which are an integral part of human interaction. At present, it remains a challenge for agent designers and system architects to produce agents that can exhibit behaviour at such a refined level. Like the agent interface constraint, this puts another limit on having agents that are nearly human.

On learners' characteristics

There is a possibility that one's visual acuity is stronger that his or her auditory capacity and vice versa. The difference may be due to one's training or past experience. Therefore the development of one's sensory preference may have an effect on the way he or she optimizes his or her visual and auditory channels in his or her cognitive mind. This exemplifies the complication in the visual-auditory operating process as explicated by the CTMLAPA model. The implication of this proposition is that sensory preference should be taken into consideration with presentation modality to get a more balanced visual-auditory interaction.

In all, it can be seen that designing a good APA learning environment is not as straightforward as it may appear. It involves the proper consideration of the APA characteristics, APA presentation and learners' characteristics as portrayed in the APL model (refer to Fig 1). An APA, regardless of look, needs to be situated in a context to actualize its social property, as a result, it needs to have coherent behaviour exhibited by the right body language, and the body language must invoke pedagogical benefits to complement the learning. All these have to be done with respect to the right type of audience whose choice of sensory preference must be congruent with the modality of the presentation. The intricate interrelationship engendered by these factors probably explains why APA research cannot be done in isolated units of experimentation (Dehn & van Mulken 2000). It is also why APA research cannot attract mass attention because few definitive results have been produced. Maybe APA research should take another view, a view that relates technology to its actual use and not what it can offer. This means that APA technology should be viewed under the lens of 'affordance'.

Briefly, the term 'affordance' means 'opportunity for action' (Kirschner *et al*. 2004, p. 49) or 'the interactions between users and tools' (Wijekumar *et al*. 2006, p. 192). Affordance offers to look at technology from the user's perspective. In the paradigm of affordance, technology is just a tool that offers opportunities for the user to interact with a specific set of features belonging to the technology to fulfil a certain task. The key is that the technology only offers the opportunity but will not necessarily result in an action by the user. Whether or not the user will use the specific features of the technology and whether after using the features the technology will help to fulfil the task will depend on the user's background characteristics, such as his or her prior experience with the tool, age level, past domain knowledge, culture, and, most importantly, intention (Gaver 1991; Wijekumar *et al*. 2006). In other words, it is the user, not the technology per se, who determines if the technology is useful. Therefore, users' characteristics are considered determinants of the use of technology in the instructional design.

Although this paper represents only some research findings, the discussion appears to veer more towards the harnessing of agent technology and ways to make agents operate more like humans. The reason why so much effort was made to produce agents that work more and more like humans is probably the fact that pedagogical agents are regarded as the 'descendants' of the ITS (Gulz & Haake 2006), which emerged in the 1990s. An ITS, as its name implies, builds on algorithms that emulate human intelligence (Beck *et al*. 2004; Corbett *et al*. 1997). Most ITSs use interface to allow text input by a learner to ask questions and text or audio output to provide answers and guidance. What falls short in this system is the presence of humans in the learning process. This makes the learning experience less humanistic. The advent of APAs naturally serves to fill this deficiency, and hence much effort is devoted to building systems that have APAs that inherit theintelligence from ITSs so

that they can teach like real humans. The historical background could have propagated the great promise of agent technology especially with the support of AI and caused many to overlook the necessity of looking for alternative means, such as using agents with less adaptivity or less humanlikeness for very specific types of learning, such as storytelling, short demonstration, or enactment of a scene. Such activities as seen under the lens of affordance would appear to be more appealing to learners because the agents employed are likely to be seen as tools, not as humans, because of the lesser intelligence that is accorded to them. Tools are familiar means of support that humans use in their daily lives. For this, humans are less likely to ascribe the same high expectation to tools as to humanlike agents, and so learners under such environments are more amenable to learning with these agents.

In summary, instead of seeing APAs as holistically a human substitute, it may be worthy to analyse APAs' contribution according to their specific strengths and weaknesses in areas such as appearance, speech, and gesture combination while taking into account the learners' background characteristics when designing an agent instruction. In other words, design should begin with the process of analysing the affordances of APAs, followed by establishing a proper match between what the APAs can offer and what the users need.

References

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

Badler N. & Allbeck J. (2001) Towards behavioral consistency in animated agents. In *Deformable Avatars* (eds N. Magnenat-Thalmann & D. Thalmann), pp. 191–205. Kluwer Academic Publisher, Norwell, MA.

Baylor A.L. (2002) Expanding preservice teachers' metacognitive awareness of instructional planning through pedagogical agents. *Educational Technology, Research and Development* **50**, 5–22.

Baylor A.L. & Chang S. (2002). Pedagogical agents as scaffolds: the role of feedback timing, number of agents, and adaptive feedback. *Paper presented at the International Conference of the Learning Sciences*, October 23–26, 2002, Seattle, WA.

Baylor A.L. & Ebbers S.J. (2003, June). The pedagogical agent split-persona effect: when two agents are better than one. *Paper presented at the ED-MEDIA 2003, World Conference on Educational Multimedia, Hypermedia & Telecommucations*, June 23–28, 2003, Honolulu, HI.

Baylor A.L. & Ryu J. (2003) Does the presence of image and animation enhance pedagogical agent persona? *Journal of Educational Computing Research* **28**, 373–394.

Baylor A.L., Kim S., Son C. & Lee M. (2005) Designing effective nonverbal communication for pedagogical agents. In *Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology* (eds C.K. Looi, G. McCalla, B. Bredeweg & J. Breuker), pp. 744–746. IOS Press, Amsterdam.

Beck J., Stern M. & Haugsjja E. (2004). Application of AI in education. *The ACM Student Magazine* [Electronic version]. Available at http://www.acm.org/crossroads/xrds3-1/aied.html (last accessed 28 May 2008)

Beun R.J., de Vos E. & Witteman C. (2003). Embodied conversational agents: effects on memory performance and anthropomorphisation. In *Proceedings of the International Conference on Intelligent Virtual Agents 2003*, pp. 315–319. Springer-Verlag, Berlin.

Buisine S. & Martin J.C. (2007) The effects of speech-gesture cooperation in animated agents' behavior in multimedia presentations. *Interacting with Computers* **19**, 484–493.

Buisine S., Abrilian S. & Martin J.C. (2004). Evaluation of multimodal behaviour of agents. A report for NIEC Project, pp. 1–22 [Electronic version]. Available at http://www.niceproject.com/publications/buisine-et-al_0311.pdf (last accessed 29 November 2004).

Chang V. & Fisher D. (2003) The validation and application of a new learning environment instrument for online learning in higher education. In *Technology-rich Learning Environments: A Future Perspective* (eds M.S. Khine & D. Fisher), pp. 1–20 World Scientific Publishing, Singapore.

Chou C.Y., Chan T.W. & Lin C.J. (2003) Redefining the learning companion: the past, present, and future of education agents. *Computers & Education* **40**, 255–269.

Clarebout G., Elen J., Johnson W.L. & Shaw E. (2002) Animated pedagogical agents: an opportunity to be grasped? *Journal of Educational Multimedia and Hypermedia* **11**, 267–286.

Clark R.C. (2005) Multimedia learning in e-courses. In *The Cambridge Handbook of Multimedia Learning* (ed. R.E. Mayer), pp. 589–616. Cambridge University Press, New York, NY.

Clark R.E. & Choi S. (2005) Five design principles for experiments on the effects of animated pedagogical agents. *Journal of Educational Computing Research* **32**, 209–225.

Corbett A., Koedinger K. & Anderson J.R. (1997) Intelligent Tutoring Systems. In *Handbook of Human-Computer Interaction* (eds M.G. Halander, T.K. Landauer & P.V. Prabhu), pp. 849–874. Elsevier, Amsterdam.

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

Dehn D. & van Mulken S. (2000) The impact of animated interface agents: a review of empirical research. *International Journal of Human Computer Studies* **52**, 1–22.

Dos Santos C.T., Frozza R., Dhamer A. & Gaspary L.P. (2002) DORIS – Pedagogical agent in intelligent tutoring systems. In *Intelligent Tutoring Systems (ITS) 2002, Lecture Notes in Computer Science Series, 2363* (eds S.A. Cerri, G. Gouarderes & F. Paraguacu), pp. 91–104. Springer-Verlag, Berlin.

Dunn R. & Griggs S. (1998) Learning styles: link between teaching and learning. In *Learning Styles and the Nursing Profession* (eds R. Dunn & S. Griggs), pp. 11–23. NLN Press, New York, NY.

Eggen P. & Kauchak D. (2001) *Educational Psychology: Windows on Classrooms*. Prentice Hall, Upper Saddle River, NJ.

Gaver W. (1991) Technology affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching through Technology*, pp. 79–84, Los Angeles, CA.

Gulz A. & Haake M. (2006) Design of animated pedagogical agents – a look at their look. *International Journal of Human Computer Studies* **64**, 322–339.

Hapeshi K. (1993) Design guidelines for using speech in interactive multimedia systems. In *Interactive Speech Technology: Human Factors Issues in the Application of Speech Input/Output to Computers* (eds C. Baber & J.M. Noyes), pp. 177–188. Taylor & Francis, London.

Hetherington E.M. & Parke R.D. (1993). *Child Psychology: A Contemporary Viewpoint*. McGraw-Hill, New York, NY.

Hietala P. & Niemirepo T. (1998). Multiple artificial teachers: how do learners cope with a multi-agent learning environment? *Paper presented at the Workshop on Current Trends and Applications of Artificial Intelligence in Education at the Fourth World Congress on Expert Systems*, March 16–20, 1998, Mexico City, Mexico.

Johnson W.L., Rickel J.W. & Lester J.C. (2000) Animated pedagogical agents: face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligent in Education* **11**, 47–78.

Kirschner P., Strijbos J.W., Kreijns K. & Beers P.J. (2004) Designing electronic collaborative learning environments. *Educational Technology Research and Development* **52**, 47–66.

Kohar H. & Ginn I. (1997) Mediators: guides through online TV services. In *Proceedings of Demo Session in CHI'97 International Conference on Human Factors in Computing Systems*, pp. 38–39. ACM Press, New York, NY.

Krenn B., Pirker H., Grice M., Baumann S., Piwek P., Deemter K., Schroder M., Klesen M. & Gstrein E. (2002) Generation of multimodal dialogue for net environment. In *KONVENS 2002, Deutsches Forschungszentrum fuer Kuenstliche Intelligenz (DFKI)* (ed. S. Basemann), pp. 91–98. German Research Centre for AI (DFKI), Saarbruecken, Germany.

Learning & Skills Research Centre, L.S.R.C (2004). *Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review*. LSRC, London.

Lester J.C., Converse S.A., Kahler S.E., Barlow S.T., Stone B.A. & Bhogal R.S. (1997). The persona effect: affective impact of aminated pedagogical agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 359–366, Atlanta, GA.

Lester J.C., Towns S.G., Callaway C.B., Voerman J.L. & FitzGerald P.J. (2000) Deictic and emotive communication in animated pedagogical agents. In *Embodied Conversational Agents* (eds. J. Cassell, J. Sullivan, S. Prevost, & E. Churchill), pp. 123–154. MIT Press, Boston, MA.

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

Lusk M.M. & Atkinson R. (2007) Animated pedagogical agents: does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology* **21**, 747–764.

McCloud S. (1993). *Understanding Comics*. Harper Perennial, New York, NY.

Magennis S. & Farrell A. (2007). Teaching and learning activities: expanding the repertoire to support student learning. Available at http://www.aishe.org/readings/2005-1/magennis.pdf (last accessed 29 July 2007)

Mayer R.E. (2001) *Multimedia Learning*. Cambridge University Press, Cambridge, UK.

Mayer R.E. (2005a) Introduction to multimedia learning. In *The Cambridge Handbook of Multimedia Learning* (ed. R.E. Mayer), pp. 1–16. Cambridge University Press, New York, NY.

Mayer R.E. (2005b) Principles of multimedia learning based on social cues: personalization, voice, and image principles. In *The Cambridge Handbook of Multimedia Learning* (ed. R.E. Mayer), pp. 201–212. Cambridge University Press, New York, NY.

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

Moreno R. (2005) Multimedia learning with animated pedagogical agents. In *The Cambridge Handbook of Multimedia*

Learning (ed. R.E. Mayer), pp. 507–523. Cambridge University Press, New York, NY.

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., 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.

Moundridou M. & Virvou M. (2002) Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning* **18**, 253–161.

Mulken S.V., Andre E. & Muller J. (1998, September). The persona effect: how substantial is it? In *Proceedings of HCI on People and Computers XIII*, London, pp. 53–66.

Nass C., Isbister K. & Lee E.J. (2000) Truth is beauty: researching embodied conversational agents. In *Embodied Conversational Agents* (eds J. Cassell, J. Sullivan, S. Prevost & E. Churchill), pp. 374–402. MIT Press, Boston, MA.

Nunes M.A.S., Dihl L.L., Fraga L.M., Woszezenki C.R., Oliveira L., Francisco D.J., Machado G.J.C., Nogueira C.R.D. & Notargiacomo M. (2002) Animated pedagogical agent in the intelligent virtual teaching environment. *Interactive Educational Multimedia* **4**, 53–61

OECD (2002). *Understanding the Brain – Towards a New Learning Science*. OECD, Paris.

Ormrod J.E. (1995) *Human Learning*. Prentice Hall, Englewood Cliffs, NJ.

Pease A. (1987) *Body*: *How to Read Others, Language, Thoughts by Their Gestures*. Sudha Publications, London.

Piwek P. (2003) The effect of gestures on the perception of a dialogue between two embodied conversational agents: a pilot study [Electronic version]. NECA Report no. IST-2000-28580. University of Brighton, Brighton.

Prendinger H., Mayer S., Mori J. & Ishizuka M. (2003) Persona effect revisited: using bio-signals to measure and reflect the impact of character-based interfaces. Fourth International Working Conference on Intelligent Virtual Agents (IVA-03), Springer LNAI2792, pp. 283–291, Springer Verlag, Kloster Irsee, Germany.

Reeves B. & Nass C. (1996) *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, New York, NY.

Rickel J. & Johnson L. (2000) Task-oriented collaboration with embodied agents in virtual worlds. In *Embodied Conversational Agents* (eds. J. Cassell, J. Sullivan, S. Prevost, & E. Churchill), pp. 95–122. MIT Press, Boston, MA.

Rose C. & Nicholl M.J. (1997). *Accelerated Learning for the 21st Century: The Six Step Plan to Unlock Your Master-Mind*, Delacorte Press, Broadway New York, NY.

Ruttkay Z., Dormann C. & Noot H. (2004) Embodied conversational agents on a common ground: a framework for design and evaluation. In *From Brows to Trust* (eds Z. Ruttkay & C. Pelachaud), pp. 27–66. Kluwer Academic Publishers, Amsterdam.

Schnotz W. (2005) An integrated model of text and picture comprehension. In *The Cambridge Handbook of Multimedia Learning* (ed. R.E. Mayer), pp. 49–69. Cambridge University Press, New York, NY.

Schunk D.H. (2000) *Learning Theories: An Educational Perspective*. Prentice Hall, Upper Saddle River, NJ.

Smith C. (2007). Sensory learning styles (Information and techniques for coaches). Available at http://www.grapplearts.com/Learning-Styles-in-Grappling.htm (last accessed 15 July 2007).

Sweller J. (2002). Visualisation and instructional design. In *Proceedings of the International Workshop on Dynamic Visualisations and Learning* (ed. R. Ploetzner), pp. 1501–1510. Knowledge Media Research Centre, Tubingen, Germany.

Walker J., Sproull L. & Subramani R. (1994) Using a human face in an interface. In *Proceedings of the CHI 1994, Conference on Human Factors in Computing Systems (CHI'94)* (eds B. Adelson, S. Dumais & J. Olson), pp. 85–91. ACM Press, New York, NY.

Wang N., Johnson W.L., Mayer R.E., Rizzo P., Shaw E. & Collins H. (2008) The politeness effect: pedagogical agents and learning outcomes. *International Journal of Human-Computer Studies* **66**, 98–112.

Welch R., Blackmon T., Liu A., Mellers B. & Stark L. (1996) The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence: Teleoperators and Virtual Environments* **5**, 263–273.

White B., Frederiksen J., Frederiksen T., Eslinger E., Loper S. & Collins A. (2002, October) Inquiry island: affordances of a multi-agent environment for scientific inquiry and reflective learning. Paper presented at the Fifth International Conference of the Learning Sciences (ICLS), Seattle, WA.

Wickens C.D. (1992) *Engineering Psychology and Human Performance*. Harper Collins, New York, NY.

Wijekumar K.J., Meyer B.J.F., Wagoner D. & Ferguson L. (2006) Technology affordances the 'real story' in research with K-12 and undergraduate learners. *British Journal of Educational Technology* **37**, 191–209.

Wissick, C. (2002) Book and software review, associate editor column. *Journal of Special Education Technology*, **17**.

Available at http://jset.unlv.edu/17.4/tasseds/wissick.html (last accessed 21 May 2008).

Wong A.F.L. & Cheung W.S. (2003) Using IT for presentation – use and misue. In *Teaching and Learning with Technology: Theory and Practice* (eds S.C. Tan & A.F.L. Wong), pp. 143–163. Prentice Hall, Singapore.

Woo H.L. & Wang Q. (2005) Designing an agent-based learning environment – a preliminary study on a pre-service teacher course. New horizons in education. *Journal of Education* **52**, 39–52.