Design of animated pedagogical agents—A look at their look

Agneta Gulza, Magnus Haakeb,∗

aDivision of Cognitive Science, Lund University, Kungs huste t, Lundagård, 222 22 Lund, Sweden
bDepartment of Design Sciences, Lund University, P.O. Box 118, 221 00 Lund, Sweden

Received 2 April 2003; received in revised form 22 July 2005; accepted 4 August 2005
Available online 26 September 2005
Communicated by E. Motta

Abstract

A well-established effect of animated agents in educational and other contexts is their potential to motivate and engage. “Increased motivation in users” is also one of the more frequent answers given to the question, “What is gained by adding an animated pedagogical agent to an intelligent tutoring system?”

To further develop and exploit this potential, there are, however, several issues that need to be resolved. In this article we discuss the visual form and look of animated pedagogical agents. A survey is presented of how the area is approached (and, in particular, not approached) in research on animated pedagogical agents. Two possible reasons are proposed as to why visual form and look are so little addressed are also proposed. We also propose and discuss some key aspects of look that merit a systematic approach in future research.

The main thesis of the paper is that users’ visuo-aesthetic experience of animated pedagogical agents is too important with respect to the goals to motivate and engage, to be treated as a secondary issue. We do not deny that there are other pressing and fundamental issues that need to be solved, such as those concerning the content of the support and the competence level of agents, as well as various design elements that can contribute to making animated agents lifelike. But we argue that visual rendering issues are pressing and need to be seriously addressed as well.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Animated pedagogical agents; Motivational issues; Visual experience; Visual form; Design space

1. Introduction

During the past decade the socio-cultural concept of learning, as interactive and collaborative processes in social actors, has had an increasing impact. This view of learning as a fundamentally social phenomenon is also reflected in the domain of computer assisted learning. There is the rapidly growing area of computer supported cooperative learning, with different kinds of support for discussion groups, on-line debates, chats, forums, arenas, etc. (Koschman, 1996; Dillenbourg, 1999). These approaches may be called extrinsically social or function oriented. A technology is provided, but the social interaction is entirely up to the human users; they must create the social context themselves. The computer systems as such can be compared to non-populated arenas, providing facilities to use and inviting and supporting those who chose to enter the arena.

The view of learning as a social phenomenon is also evidenced in approaches in which an already populated social arena is created by technology—intrinsically social or content oriented approaches. Examples are scenario-based systems inhabited by social characters (Schank and Neaman, 2001) and systems using animated pedagogical agents, which is the focus of this article (e.g. André et al., 1998; Cassell et al., 1994; Paiva and Machado, 1998; Cassell and Thónisson, 1999; Shaw et al., 1999; Johnson et al., 2000; Lester et al., 2000). In such systems various social attributes, such as the abilities to express socially appropriate behaviours and to handle them, are implemented in
the computer technology. A designed social context is provided, and can be entered and further developed by several users, but also by a solitary user.¹

1.1. Animated pedagogical agents

An animated pedagogical agent can be considered an extension of an intelligent tutoring system (Shaw et al., 1999). Work on intelligent tutoring systems goes back to the early 70s (Laurillard, 1993). The intention is to provide students with an individualized tutor through the use of artificial intelligence. With the three components of domain knowledge, student model and teaching knowledge, the system should be able to survey a student’s actions and progress, provide feedback and give contextual advice and support for problem solving.

But whereas a classic intelligent tutoring system—as manifested only in textual output—is invisible and fairly abstract, the addition of an animated pedagogical agent to the interface provides elements of embodiment, visibility and personality. In addition to the ability to communicate in an intelligent manner, a pedagogical agent should, according to Lester et al. (2001), have socio-emotive abilities and be lifelike. It should be visually present, by means of gestures, facial expressions and so on, and have a rich and interesting personality Lester et al., 2001. Consequently, the addition of animated agents to intelligent tutoring systems opens up the possibility for learners to have a personal relationship and an emotional connection with the agent, which in turn may promote interest in the learning task (Moreno et al., 2001).

The addition of animated agents to intelligent tutoring systems can, in other words, be seen as an attempt to fulfil the need for a social context for learning in these systems (Kearsley, 1993).

Several educational systems and prototypes that incorporate various kinds of animated pedagogical agents exist today.² In some systems the animated agent primarily acts as teacher, instructor or coach—e.g. AutoTutor (McCaughey et al., 1998); Vincent (Paiva and Machado, 1998); Whizlow (Lester et al., 1999); Adele (Shaw et al., 1999); Cosmo (Lester et al., 2000); Spin the Dolphin (Oivatt and Adams, 2000); Herman the Bug (Lester et al., 1997); Talking Head (Moundridou and Virvou, 2002); Laura (Bickmore, 2003). In other systems the animated agent primarily acts as a learning companion—e.g. Trouble Maker (Aimeur et al., 1997) and Steve³ (Richell and Johnson, 2000).⁴ In still others the animated agent primarily acts as a presenter or guide—e.g. Olga (Beskow and McGlashan, 1997); Jack (Noma and Badler, 1997); Rea (Cassell et al., 2000); Will (Churchill et al., 2000). Furthermore, some systems are inhabited by several agents, such as the PPP persona presentation team (André et al., 1998) and Steve, with multiple instructor agents and team-mate agents (Rickel and Johnson, 2000) (Fig. 1).

The objection might be raised that presenter or guide agents are not properly educational. Indeed the agents in the presenter or guide agents group above are, in general, not implemented in educational systems and do not focus on educational dialogue and interaction. Yet given that presentation of material and guidance through material are teaching elements, the boarders are not clear-cut.⁵ In this article we regard learning and pedagogy in a broader sense and thus include animated agents that have more of a presentational and guiding role.

1.2. The potential to motivate and engage

The most well established effect of animated agents, in educational as well as in other contexts, is their potential to make the experience of a program more engaging (Walker et al., 1994; Takeuchi and Naito, 1995; Kodai and Maes, 1996; Lester et al., 1997; André et al., 1998; van Mulken et al., 1998; Dehn and van Mulken, 2000; Rickenberg and Reeves, 2000; Moundridou and Virvou, 2002). Engagingness is primarily measured by means of interviews or questionnaires, where subjects relate their experiences and attitudes: to what extent did they appreciate or enjoy the experience; to what extent did they feel involved in the interaction (Dehn and van Mulken, 2000). In addition there are related behavioural measures: how long does a user stay on in a learning environment; how much activity does she or he exhibit; how willing is he or she to use the program again, etc. The relation between attitude measures and behavioural measures is an effect of the relation between intrinsic motivation and activation (Malone, 1981; Keller, 1983). If a program is found engaging—that is, experienced as involving, interesting or as having impact—it is likely that users will become more active, stay on longer, and produce more. Engagingness in this sense is not to be equated with entertainment.⁶ To be entertained by an interface agent does, indeed, imply an important form of engagement, but one can be engaged on several other grounds as well. The impact that an engaging program has does not even have to be of a pleasant nature. Walker et al. (1994), using two different animated agents, one with a stern facial expression and one with a neutral, demonstrated that subjects liked the version with the stern

¹There are also extrinsically and intrinsically mixed social forms, as for instance in using a teachable agent system or a social scenario-based program via a computer supported co-operative learning system or in a pedagogical agent system that supports multi-user collaborative exercises (Shaw et al., 1998).

²The reader is also referred to Johnson et al. (2000), in which a number of animated pedagogical agents are described and discussed.

³That is, Steve as a virtual team-mate; the system also incorporates Steve in an instructor role.

⁴There is also the set of visually designed and personified learning companions in EduAgent (Hietala and Niemirepo, 1998). These agents are, however, not animated.

⁵As an illustration of the vagueness, Gandalf (Cassell and Thörisson, 1999) is classified as “a pedagogical assistant” in Dehn and van Mulken (2000), whereas Johnson et al. (2000) state that Gandalf “does not address tutorial dialogue”.

face less, but spent more time, wrote more comments and made fewer mistakes with this version.

An increase in engagement may, in turn, have effects on learning achievement. It might lead a student to interact with a system more frequently or increase the time spent within a learning environment, which may result in superior learning achievement in terms of taking in more, understanding more or remembering more (Dehn and van Mulken, 2000; Lester et al., 2001).

Schank and Neaman (2001) point out three ways in which motivation may affect cognition. One is the participation issue: motivation may affect whether a student participates at all in a learning activity. Another is the indexing issue: motivation affects the way that memories are organized. A third is the attention issue: motivation affects the quality of attention during the learning experience, which in turn affects what is remembered. The attention issue is complex, as several authors writing on animated agents acknowledge. Too little engagement involves a low quality of attention and poor memory. On the other hand, there is the risk that an animated agent may be engaging to the point that it attracts attention in such a way or to such an extent that it functions as a distractor from what should be learned. (Cf. van Mulken et al., 1998; Rickenberg and Reeves, 2000; Moreno et al., 2001.)

A central motive for the development of animated pedagogical agents is to enhance learning in students. Given the strong, albeit complex, relationship between motivation and learning, the potential of animated pedagogical agents to engage and motivate is relevant in this regard. This potential is the focus of this paper: specifically we address issues of the visual form of pedagogical agents, a hitherto neglected aspect that we argue is of great importance for motivation and engagement.

1.3. The structure of the paper

The paper is organized as follows: In the following, second, section we survey issues with relevance to the motivational effects of pedagogical agents that have been studied. The issues are divided into two categories: (1) user groups and study domains (2) elements of animated agents such as personality, linguistic capabilities, gestural capabilities, and so on. It is within the second category that we point out the visual rendering or look of agents as a neglected design element.

---

Fig. 1. The agents Laura (upper left), Cosmo (upper right), Rea (lower left), and Steve (lower right). Laura: © Affective Computing Group, MIT Media Lab; Cosmo: © IntelliMedia, North Carolina State University; Rea: © GNL Group, MIT Media Lab; Steve: © CARTE, University of Southern California.
In the third section we survey the literature on how visual rendering is (not) treated in research on animated pedagogical agents—in particular in the context of motivational effects. This is followed, in the fourth section, by a proposal of two possible reasons as to why visual rendering is not included in research. Thereafter, in section five and six, we argue for the importance of addressing visual rendering issues in future research on animated pedagogical agents. Four key aspects of look that merit to be systematically studied are identified: realism versus iconicity; face; body and costume design; visual style. The seventh and final section concludes by presenting a broader view of the issues discussed in the paper.

The main thesis of the paper is that the visuo-aesthetic experience of animated pedagogical agents is too important with respect to motivation to be treated as a secondary issue. We do not deny that there are other pressing and fundamental issues that need to be solved, not the least on the intelligent tutoring system level regarding the content of the support and the competence level of agents. But we will argue that visual rendering issues are pressing and fundamental as well and thus ought to be seriously approached.

2. Motivational effects of animated pedagogical agents—a survey of studied issues

2.1. User groups and study domains

With the development of the research on animated agents, the questions posed have become more modulated. Instead of asking whether or not animated agents have a certain effect—such as being engaging and motivating—it is being asked for which people, in which conditions and for what kinds of domains that a certain effect can be shown (cf. Lester et al., 1997; van Mulken et al., 1998; Rickenberg and Reeves, 2000).

Several studies have demonstrated that users react differently to an animated agent based on their own personality and other dispositional traits. In particular the engagingness effect varies. Reeves and Nass (1996) and Nass et al. (2000) showed that users like agents that match their own personality on the introversion/extroversion dimension more than agents which do not. Rickenberg and Reeves (2000) demonstrated that animated agents affected arousal reactions of users differentially as a function of whether users tended towards internal or external locus of control. The authors suggest that external versus internal control orientation may be part of the explanation as to why some people like and some don’t like animated interface characters. Whether arousal—which relates to the experience of engagement – is positive or negative depends, as the authors point out, upon various factors, such as the strength of the arousal. In a learning context, it is negative if a student is non-engaged or bored, but it is also negative if she or he is aroused to the point of distraction. A middle ground of arousal and engagement, on the other hand, has positive effects on attention and memory. (Compare this to the attention issue described in the last section.)

Regarding the domains for which animated agents may or may not have an engaging effect, van Mulken et al. (1998) showed the following: In a program that served to explain a technical device, subjects experienced the presence of an animated agent as more entertaining than that of just a pointing arrow. But in a program with the function of introducing fictitious employees of a research institute, the subjects’ entertainment ratings of an interface with an animated agent versus one with a pointing arrow did not show any substantial difference. The explanation put forth by the authors is that the interface in the employee task might have been entertaining from the start and that the additional agent made no difference to the user.

2.2. Design elements of animated agents

In addition to asking whether the presence of an animated pedagogical agent in a certain context has motivational effects, there is also the question of what elements of animated pedagogical agents produce these effects. Among design elements of animated pedagogical agents that have been extensively investigated are the following:

- Movement characteristics, in particular gestures, hand movements (Lester et al., 1998, 2000; Cassell et al., 2000; Massaro et al., 2000).
- Facial expressions (Lester et al., 2000; Massaro et al., 2000; Poggi and Pelachaud, 2000).
- Voice characteristics (Nass et al., 1994; Nass and Gong, 1999; André et al., 2000).
- Dialogue and conversational characteristics (Cassell et al., 1998; McCauley et al., 1998).
- Emotional expression via voice, gestures, facial expressions (Bècheiraz and Thalmann, 1998; Badler et al., 2000; Ball and Breese, 2000; Lester et al., 2000; Massaro et al., 2000).
- Personality realized via voice, gestures, facial expression, verbal communication (Ishibster and Nass, 1998; Taylor et al., 1998; André et al., 2000; Ball and Breese, 2000; Churchill et al., 2000).

A design element, however, that is more seldom examined is that of look or visual form. This state is the focus of this paper, and in the next section we take a look at how look is (not) treated in the research literature.

---

8We are not discussing design features related to the intelligent tutoring system that underlies the agent.
3. A look at look in the literature on animated pedagogical agents

3.1. Personality

Dryer (1999) explored the effect that some elements of visual rendering have on how users perceive the personality of agents. He presented subjects with a set of animated characters to measure their perception of the characters’ personalities, and found that characters perceived as extroverted and agreeable tended to be represented by rounder shapes, bigger faces, and happier expressions, while characters perceived as extroverted and disagreeable were typically represented through bold colours, big bodies, and erect postures.

André et al. (2000) also explored look in relation to perception of the personality dimensions: disagreeable, agreeable, extrovert and introvert. Specifically they studied effects of attempted mismatches between a character’s look and voice with respect to the four personality dimensions. Subjects’ comments suggested, according to the authors, that both the look and voice of a character are important cues to its personality and interest profile. Therefore, the authors propose, the possibilities of reusing the look and voice of characters for different roles are limited.

In research on animated agents, personality indeed is a key issue and concept. And focusing on engagingness of animated agents, much indicates that personality is central (Isbister and Nass, 1998). However, whereas there is a bulk of research on how animated agent personalities can be expressed in verbal communication, by gestures, by voice and by facial expressions, there is comparatively little on visual rendering, in spite of it being well established that personality is not expressed by communication alone, not even if one includes non-verbal communication such as gestures, voice and facial expressions. We will return to the issue of personality in the section on future research.

3.2. Facial expressions

Another key issue in research on animated pedagogical agents is facial expressions. Notably, there is almost no research that involves systematic studies of different facial looks, i.e. the underlying facial form on which the expressions are generated is not subject for analysis. The same holds for movements and gestures, where the underlying body—and costume—on which gestures and movements are superimposed is little articulated. This may be compared to the design process in animated movies, where often an extended period of time is dedicated to finding the right face, body and costume, before the animation work begins.10

The main research focus in studies on faces of animated agents is on different facial expressions based on one look, and user experiences of those (Walker et al., 1994; Massaro et al., 2000). A frequently cited study is that of Walker et al., 1994, who studied how facial expressions affected users’ experiences and performance. The visual form was held constant in the sense that one relatively stern expression and one relatively pleasant (neutral) one were derived from the same underlying image (Fig. 2). (As mentioned previously, the stern face created a bad impression but was good for productivity.) Thus, the visual form and style of faces is not included in Walker et al.’s, 1994, empirical explorations. However, the authors do acknowledge the issue in posing some important questions: “Should human facial realism be a goal? If so, whose face would appear?” Furthermore, they mention that it is established that attractive faces improve people’s response to advertisements.

Another focus is on comparing face conditions with non-face conditions (Koda, 1996; Parise et al., 1996; Sproull et al., 1997). In these studies the focus is, in other words, on symbol versus non-symbol—not on the different visual renderings of the symbol.

One study that directly addresses an aspect of facial look is that of Lee and Nass (1998) who explore the role of visually represented ethnicity of animated agents. In the study it was demonstrated that people perceived agents with similar ethnicity to their own to be more like themselves, more attractive, more trustworthy and more persuasive. A general conclusion of the study is, according to Nass et al. (2000), that “appearance is a critical component of how people access agents and for their preferences in terms of looking at and even interacting with”, and that appearance influences peoples’ cognitive assessments.

3.3. Appearance

Lee and Nass (1998) and Nass et al. (2000), thus, include visual form and look in the broader term appearance. In general, however, when the issue of appearance is examined in empirical investigations on animated agents, it is movements, gestures, postures, facial expressions, etc. that are being explored (Isbister and Nass, 1998; Massaro et al., 1998). The effects of varying these aspects are studied, while look or visual form is held constant. The particular visual form, that is thus held constant, is, in general, unproblematized and seemingly randomly chosen.

Churchill et al. (1998, 2000) claim that the design of the character’s appearance is central to charting out what design issues ought to be explored for situated design issues...
conversational characters.\textsuperscript{11} The authors explicitly mention the design of the Look\textsuperscript{12} and propose five dimensions of look to be explored: the degree of humanoidness, the degree or stableness versus change-ability in appearance (e.g. morphing), the degree of animation versus staticity (i.e. the extent to which the character can move), 2D or 3D visual rendering and degree of realism—from photorealism to line drawing. The authors point out that not much is known on how these dimensions affect human user reactions, and in particular degrees of realism. The acknowledgement of visual rendering issues is thus very explicit in Cassell et al., 2000. Look and visual rendering are included, and dimensions listed in the research agenda they put forth. Yet in the actual development of their own character—Will—no exploration of these design spaces seems to have taken place. The visual form, the look of Will is more or less left without discussion.

In summary, empirical exploration on aspects of visual rendering of animated agents is sparse, even among researchers who acknowledge its importance.

\textsuperscript{11}Situated or embodied conversational agents are “specifically conversational in their behaviour and specifically humanlike in the way they use their bodies in conversation” (Cassell et al., 2000).

\textsuperscript{12}In (2000) the authors replaced the term ‘look’ with the term ‘embodiment’.

4. Two possible explanations for the neglect of visual rendering issues

One possible explanation as to why visual rendering is neglected in research on animated agents is that it cannot be readily approached with existing research methodologies. Much of the research on animated agents is technologically driven, and computational approaches have a large impact. Work on gestures, facial expressions, voices, etc. largely involves the development and refinement of computational algorithms. Once analyzed, voices or gestures may be produced and reproduced by such algorithms. It seems unlikely, though, that computational algorithms could handle the process of selecting or deciding upon a look. Even if the visual rendering of one agent is analyzed, it is hard to imagine a program that will successfully generate the visual appearance for the animated agent in another educational program. Compare, for instance, voice and appearance. The perception of what is a pleasant voice is likely to be more stable over time than the perception of what is a pleasant appearance, in terms of clothing, body shape, hair, etc. The latter is more sensitive to trends. Thus complementary means of handling such aspects seem required.

It should be observed that some of the research on gestures and facial expressions in animated agents does not

Fig. 2. The two faces used in the study of Walker et al. (1994). Left: pleasant/neutral expression; right: stern expression (compare the eyebrows). © Walker et al. (1994) & Association for Computing Machinery (ACM).
involve computation. For instance, some of the research on design of gestures for agents is based on ethnographic and sociological research (O’Neill-Brown, 1997; Cassell, 2000; de Rosis et al., 2004). Nevertheless, most of the work on the design of gestures in agents strives for computational algorithms, and the same goes for work on facial expressions and voices. But there is hardly any corresponding work to be found on visual rendering.

Another possible explanation for the neglect of the look of animated agents may be that the influence of look on emotional and intellectual processes is not readily accepted, although empirically well established. Plenty of research supports the halo effect, i.e. the commonly held view that good-looking people have other positive traits such as being independent, sociable, intellectually capable and interesting (Brigham, 1980; Langlois et al., 2000). Unattractive people are considered as less socially competent, less willing to cooperate, more dishonest, unintelligent, psychologically unstable and antisocial. (Jones et al., 1978; Mulford, et al., 1998; Langlois, et al., 2000). Negative reactions to unattractive people are also more severe. For instance, Ahola (forthcoming) presented subjects with veridical crime descriptions and photos of people that allegedly had committed certain crimes. The study demonstrated that subjects recommended longer prison sentences for people who were not good-looking than for people who were. Furthermore the crimes were perceived as less serious. Pre-studies had been assessed (see section 4).

Tensions of appearance and observable physical cues profoundly affect both beliefs and behaviour (Milord, 1978; Schneider et al., 1979; McArthur, 1982; Kalick, 1988).

Nass et al. (1994) and Sproull et al. (1997) have shown that the same, in many cases, holds for animated agents. The Lee and Nass study (1998), mentioned earlier, demonstrates that aspects of appearance are critical for how people access agents, and for their preferences in terms of looking at and even interacting with agents.

A possible explanation for the role of visual appearance is that it is significant for the representation of personality. Branham (2001) borrows the drama theory term physical personality of a character to refer to the aspects of appearance, which immediately and with no acquaintance, produce an impression of personality, and which initiate a set of attitudes and expectations. Among those aspects are many visual aspects such as shape, height, sex, race, physical attractiveness, hair, clothing, makeup, facial shape, facial hair, and so on. In contrast to how the wealth of such visual cues is immediately picked up in the encounter with another human being, the slow linear stream of spoken information is incredibly small (Gard, 2000). Likewise Berscheid and Walster (1974) note that, “our appearance telegraphs more information about us than we would care to reveal on a battery of personality inventories […] From flame-coloured hair through flat feet, few aspects of appearance fail to provide kernels of folk insight into another’s nature.” Regardless of how accurate the ideas are, that we thus build, we do build them. Toby Gard, designer and lead artist of the game Tomb Raider and its main character, Lara Croft, says in discussing the design of computer game characters, that “a person’s first impression of a character will almost certainly come not from what they do, think, or say, but what they look like. If the character makes a good first visual impression, players will likely stay focused on it, allowing you to further entice them with the character’s personality.” (Gard, 2000, p. 4–5). Likewise, Lasseter (1987), in a paper on animation principles, points out that look is linked to the animation term appeal, where appeal is “[a]nything that a person likes to see [and that the] eye is drawn to […] Where the live action actor has charisma, the animated character has appeal.” (Lasseter, 1987, p. 42).

Furthermore, research lends credit to the folk psychological notion that “first impressions are lasting”. Impressions of someone’s personality based on physical appearance, not only persist but also deepen over time (Mathes, 1975). Given that people seem to treat animated agents in ways similar to how they treat human beings (Reeves and Nass, 1996), it is likely that the principle “first impressions are lasting” holds for animated agents as well. Gard (2000) claims in a text on computer game characters, that even though our opinions on a person’s personality may be reformed after a while, it will “for a long time […]

5. The importance of visual rendering of animated pedagogical agents

In summary, compared to the amount of research on other design aspects—facial expressions, gestures, dialogue characteristics, etc.—there is little that has been carried out on the visual aspects of pedagogical animated agents, that is, on different faces, bodies and clothing; on degrees of iconicity in the visualization; and on visual styles. With respect to the pedagogically central potential of animated agents to be engaging—to increase involvement and contribute to the impact of learning activities—this research gap may be unfortunate. When people interact with real people, there is ample evidence that interpreta-

---

13One exception is the work of Branham (2001) that attempts to modelling trait impressions of faces, such as facial maturity, and to have agents that dynamically can generate faces representing specific physical personalities. She expresses, however, the reservation that altering the physical personality of an agent in this way could lead to user confusion.

14These evaluations had been assessed in pre-studies.

15Swedish proverb. (Döm inte hunden efter haåren!)
still [be] filtered through our preconceptions based on our first impressions. So to create a really good character, you have to control all of the visual clues that people use to judge each other and establish a clear, unified message to make players interested in—and ultimately like—your character." (Gard, 2000, p. 3), see Fig. 3.

An implication of all of this is that if visual appearance and look are just left to happen, rather than being carefully considered and articulated in research and development on animated pedagogical agents, one risks ending up with agents that fall short of motivating, engaging and adequately impacting users. An illustrative example of what can happen is reported by de Rosis et al. (2004). An animated character was designed for a natural-language interface for a legal information system in Italy. Initially the character was designed as a very attractive young female assistant, since the developers assumed that the typical user of their system was going to be a male lawyer. However, after realising that, in fact, the lawyer’s (female) secretary was the one who most frequently used the system, they became aware that the appearance and behaviour of the character disturbed these users. They therefore designed a new character, with a more classical attire and a more professional communication style. The point is that it would not have been sufficient to redesign behaviour, including linguistic behaviour, facial expressions, voice and gestures. The look—the physical personality—also had to be redesigned.

6. Future research on visual rendering of animated pedagogical agents—four key aspects

There are four key aspects of look that we believe merit systematic exploration: realism versus iconicity; face; body and costume design, and visual style. This is not an exhaustive list. The research space may well be expanded and redrawn, but according to the survey we have conducted there is reason to look at least into these four aspects. A point of caution is needed: The impact and engagingness of an animated pedagogical agent ultimately depends upon the agent as a whole—an gestalt phenomenon including all visual aspects, together with other aspects, such as voice, dialogue, communicative style, facial expressions, and the design of the underlying intelligent tutoring system. The whole is more than the sum of its parts. Nevertheless, we need to scrutinize different aspects and attempt to understand them individually. The main argument of this paper is that the visual aspects must be included in this scientific endeavour.

On the one hand, we will in the discussion of the four aspects of look relate evidence from the area of human-human interaction. The rationale is that many interaction patterns from human-human interaction seem to recur in human-computer interaction, specifically in human-agent interaction (e.g. Reeves and Nass, 1996). On the other hand, we will look into the areas of animated film and computer games, in which animated characters are central. In both areas, there is one reference that keeps appearing, and is referred to as a sort of canon, namely the book The Illusion of Life: Disney Animation by two leading Disney animators Thomas and Johnston (1984). While the explicit focus of the book is on the art of animation and illusion of life, the importance of the underlying visual form is also highlighted throughout the book. For instance, it is stated (Thomas and Johnston, 1984, p. 222) that “[w]e must study the design carefully, questioning the shape of his whole figure, his costume, his head, cheeks, mouth, eyes, hands, legs, arms—even the setting he is in and how he relates to it. Is the scale correct? Is it drawn to give the best

Fig. 3. Lara Croft (middle)—main character of the game Tomb Raider. Left and right are two concept versions of Lara Croft illustrating the concept of a simplified, yet powerful design (left) with strong and immediate visual characteristics versus an overworked design (right) creating visual bewilderment and confusion (Gard, 2000, p. 3). Left and right: © Toby Gard; Middle: © Core Design.

17The reference is also frequently found in the literature on animated pedagogical agents.
advantage of the character? Does it support and fortify his personality so that he feels dominating or timid or clumsy or defiant, or whatever he is supposed to be? This is as much a part of the problem as the type of movements he has, the timing of them, and the acting in both body attitudes and facial expressions.”

6.1. Realism versus iconicity

The notion degree of iconicity signifies, in this text, the degree to which a depicting representation is simplified and reduced (Fig. 4). Several researchers and theoreticians have put forth ideas on realistic versus iconic agents with respect to involvement and engagement effects in users. Welch et al. (1996) argue that pictorial realism increases involvement and the sense of presence in a digital environment. The authors propose that pictorial realism may even be a condition for human cooperation with an animated agent. Nass et al. (2000) argue that each aspect of appearance of an animated agent should be as similar as possible to the user group in question, and that it is a design goal to create “embodied conversational agents that accurately mirror humans”—something that also ought to imply realism in visual rendering.18

On the other hand, McCloud (1993) in his seminal book Understanding Comics argues that audience involvement is often increased by iconization. The underlying mechanism, according to McCloud, is that the concept and image of oneself is highly iconic, in contrast to that of other people in one’s environment. When people interact, they usually look directly at one another, seeing the features of others in vivid detail. Each one also sustains a constant awareness of his or her own face, but this mental image is of an iconic nature (McCloud, 1993). Therefore, identification and social affinity with an agent come more naturally and effortlessly in response to an iconic agent. This, in turn, can increase the impact that the agent has on users (Fig. 5).

McCloud also puts forth the masking effect as an important design method. This effect implies that characters are iconized but the background is realistic. In the world of animation—where the masking effect in many cases also is a practical necessity—Disney has used it for decades with impressive results with when it comes to engagingness of characters.19

Lee and Nass (1998) report differences in user responses and interactions dependent on match or non-match of ethnicity between user and agent. In the matched condition participants perceived the agent to be more socially attractive and trustworthy. They also conformed more to the decisions of their agent partner and perceived the agent’s arguments to be better. The authors accordingly suggest that if a user group is ethnically mixed, one should provide multiple agents with correspondingly different ethnicities from which the users can choose. The agents in the Lee and Nass study are, however, implemented as full motion videos of people, that is, they are photorealistic. An alternative approach, in line with McCloud’s arguments, 18At the same time that the article seems to argue for realism in visual rendering, it brings up an experiment where very simplified agents in the shape of stick figures were used, and it is emphasized that very few cues are needed in order to elicit social attributions (cf. Nass et al., 1994).

19And it can be found in many popular comics, from Asterix to Tintin to works of Jacques Tardi. In Japanese comics and animated films, as well, the masking effect has permeated the entire genre.
might be to a more iconic agent with a minimum of ethnicity. This seems to be Disney’s recipe in recent years, with many iconized characters that are ethnically ambiguous in their visualization. At the same time, notably, believability remains a central goal in the animated film domain. Gard (2000), analysing computer game characters, relatedly holds that, “there is a vast difference between realism and believability [...] I feel you can always get a stronger, more universal emotional response from high-quality hand animation than you ever can from motion capture.” (Gard, 2000, p. 6).

In summary, we find diverging claims regarding realistic versus iconic agents with respect to their impact and ability to involve. This indicates a need for systematic studies. Learning from previous results, one variable to consider in such studies is different user groups. There may be differences between cultures and sub-cultures in responses to realistic versus iconic animated agents. Yet another variable of interest is the role of the agent, for instance, a virtual teacher versus a virtual learning companion. If McCloud’s (1993) framework is applied, a teacher character, representing the other to a higher extent than a learning companion, might benefit from more realism in the representation. A learning companion character, being to a higher extent conceived of as an extension of oneself, may, on the other hand, benefit from a more iconic representation.

6.2. Face

There is, as already noted, extensive research on facial expressions of animated agents, but an analysis and articulation of the choice of a face as such, is rarely undertaken. Yet, it is likely that the facial form as such of an agent has great impact. As Magli (1989) remarks in The Face and the Soul, a face as such is loaded with complex cultural expectations. Once someone or something—as an agent—is endowed with a face, it enters a cultural sphere and becomes a player in a social arena. In human-human-interaction, several studies have shown that a person’s face is one of the most decisive factors in relations with other people (Alley, 1988; Ahola, forthcoming). People react to others based on their facial features and make all sorts of subconscious assumptions based on looks. Furthermore, they also believe that a face indeed provides valuable clues regarding a person’s character (Liggett, 1974). Basic patterns in human-agent interaction appear to be similar to those in human-human interaction (Reeves and Nass, 1996). It is therefore likely that the face of an agent, both in realistic and more iconic visualizations, plays a central role for interaction.

Some aspects of face have been rather extensively researched with respect to human-human interaction. One concerns the halo effect, mentioned earlier in the text: that we treat attractive people better than we do ugly people, and make various subconscious assumptions based on attractiveness. Faces certainly play a central role in terms of aesthetic appeal. One would, as Branham (2001) points out, be hard-pressed to name one culture that did not in some way or another encourage its members to enhance the aesthetic appeal of the face.

Another finding is that baby-faced people (Fig. 6) are considered more naive, honest, warm and kind-hearted than others. They are also seen as weaker and more submissive as well as more helping and caring, but also more in need of protection (Branham, 2001). Mature-faced individuals, in contrast, are more likely to be perceived as experts and to command respect (Zebrowitz, 1997). It should be noted that it is not only prototypical baby faces, such as the face of a real baby, that elicit these reactions and assumptions, but also faces that resemble the prototypical case more weakly (Branham, 2001; Zebrowitz, 1997). Correspondingly, Thomas and Johnston (1984) describe how a face that is a little bent downwards, with eyes looking up at you, has connotations of innocence and vulnerability, see Fig. 6.

A third aspect found to have an impact is the facial expression of the neutral face. We are, of course, affected by dynamical facial expressions, such as smiles and frowns—but also by the expression of the face in a resting position. For instance, the lips naturally turn upwards on
some faces in resting position. Such people are viewed more positively. They are considered friendly, kind, easygoing, and non-aggressive (Branham, 2001). In a similar vein “faces that have features indicative of anger or hostility, e.g. low-lying eyebrows, thin lips, and withdrawn corners of the mouth, are perceived to be more threatening, aggressive, and dominant” (Branham, 2001, p. 3).

Furthermore, similar facial expressions and an identical verbal utterance posed may be perceived as suggesting quite different motives depending upon the face of the person who is speaking. Branham (2001), reports a scenario where people are shown a photograph of a secretary and given a written description of the dialogue between the secretary and her employer, a young man. The secretary has just been thanked for taking a dictation and asks if there is anything else her employer would like her to do. When asked what the secretary means by her question, people vary in their responses depending upon which photograph representing the secretary they are shown. “Some faces consistently suggest seductive motives, others ambitious intentions, and some a polite way of taking leave,” (Branham, 2001, p. 3).

### 6.3. Body and costume design

In computer games and animated film, both body and especially costume are considered important aspects in the communication of a character’s personality (Thomas and Johnston, 1984; Gard, 2000). Thomas and Johnston, 1984, declare that “[t]he value of the costume in creating a personality cannot be overestimated” (Thomas and Johnston, 1984, p. 415).

The visualization of the figure can be discussed both as a whole (implying a gestalt phenomenon), and with respect to separate parts such as the costume and the figure in terms of head, cheeks, mouth, eyes, hands, legs, and arms. As the costume is superimposed on a body we find reason to treat these two aspects together.

Within social psychology, there has been extensive research on stereotypes and on how people categorize each other on the basis of visual appearance. Concerning bodies, there seems to be agreement on three major body stereotypes identified by Sheldon as early as 1940 (Sheldon et al., 1940); the muscular, the fat and the thin (Fig. 7). Much research has since then verified the findings of Sheldon, showing that muscular bodies are assigned positive traits, fat bodies negative traits, and thin bodies are somewhere in between (e.g. Iwawaki and Lerner, 1976; Butler et al., 1993). In a more detailed study, muscular people were seen as being more attractive, healthy, brave, competitive, and adventurous, as well as less intelligent, more intolerant, and temperamental (Ryckman et al., 1991). In his book The Psychology of Stereotyping, Schneider (2003) reports on powerful sub-types, where couch potatoes are rated negatively, while Santa Claus types are rated more positively, both belonging to the fat body type. Surprisingly enough, not many height stereotypes are reported, although this is an obvious feature of our physical presence. It should be noted that all related stereotypes concerning the body vary depending on gender, culture, fluctuations in the market, etc. For example, Schneider (2003), mentions a report about a shifting of the ideals of feminine beauty towards a more Rubenesque type in the US as the affluent 1920s turned into the depression of the 1930s (Fallon, 1990).

Hair, cosmetics and costume are extremely important as visual clues. Hair has a long history of symbolic impact, the old Assyrian kings having impressively curled false beards and Samson losing his strength as Delilah cut it off. Through the ages, people have manipulated the style and colour of their hair, in order to adapt to different social and cultural contexts or to signal specific stereotypes. Furthermore, it has repeatedly been shown that we draw inferences about people based on the clothes they wear (e.g. Bardack and McAndrew, 1985; Kaiser, 1985; Johnson and Roach-Higgins, 1987). Again such inferences are heavily dependent upon cultural norms and contexts. For instance, norms about what women and men should wear on various jobs change rather quickly. However, an obvious rule of thumb is that culturally approved use of cosmetics and clothes is associated with positive stereotypes (Schneider, 2003).

In the world of animation, Thomas and Johnston (1984), rarely explicitly articulate issues of how to design the figure and costume. But the issues are implicitly handled between the lines and in many of the examples illustrating the step-by-step sketching and development of different figures. An example is the many pictures illustrating the visual look of the
three Good Fairies in *Sleeping Beauty*, where there is also a short comment on how some late changes in the drawing of the figures and especially their hats completed the final design (Thomas and Johnston, 1984, pp. 401–405), see Fig. 8.

In the area of computer games the focus notably shifts from the design of the body towards costume design. According to Gard (2000), discussing characters in computer games, “[t]he visual design of a character can be split broadly into two aspects: physiological form and the clothes worn (if any). Physiological differences between one human and another are fairly slight; there is some variation in skin tone, size, hair, build, and weight. Gender is the only major variance [...] Clothing, however, varies greatly in colour, shape, purpose, and significance. That’s why costume design is so important.” (Gard, 2000, p. 4). Furthermore, Gard (2000) presents some fundamental guidelines for costume design applied in the computer game industry.

As a complement to costume, Gard (2000) as well as Thomas and Johnston (1984) speak of additional elements and specific articles such as glasses for the establishment of the personality of a character. The wearing of eyeglasses has been intensively studied, and most recent research agrees that glasses are associated with mental competence and intelligence (Terry and Krantz, 1993; Hellström and Tekle, 1994), as well as diminished ratings of social competence and forcefulness (Elman, 1977; Terry and Krantz, 1993).

In addition to the discussion above of the body-costume dimension, the performing arts and film have a long tradition of addressing costuming. It is likely that a clever mix of consistency and subtle changes in the costuming of an agent can be used to affect user engagement.

### 6.4. Visual style

In contrast to face, body and costume, there is little empirical research on visual style. In this section, therefore,
we rely on experience from visual media: computer games, animated film and comics. The concept of *visual style* is illusive, but could roughly be regarded as the manner in which static and dynamic visual elements are expressed, arranged and animated, individually as well as on the whole, thereby evoking particular associations, experiences and moods.

In the area of computer games, there has been a development from text based and simple graphic systems towards games embedded in increasingly spectacular graphical packages. Likewise the new genre of computer-animated films is heavily promoted on the basis of astonishing graphical effects, where Disney-Pixar is able to excel with remarkable water effects in their latest creation *Finding Nemo* (Disney-Pixar, 2003). Of course, the eventual success of new animated films and computer games does not depend on visual effects alone. Both in animated film and computer games, we find awareness about the importance of the story and overall context. At the same time what is really crucial is the intricate *interaction* between the visual aspects and the story and context. In order to obtain believable personalities that act in a believable context, it is necessary that the characters and the story mesh well with the visual style (Thomas and Johnston, 1984; Gard, 2000).

In comics and cartoons, visual style can be discussed in terms of drawing styles where the inked line has properties such as direction, shape, and character. By means of such properties it is possible to establish a style or an overall mood in the drawings, such as: cool sophistication; whimsy and youthful innocence; depravity and morbid decay; etc. (McCloud, 1993), see Fig. 9. However, in traditional two-dimensional cell animation, the line and the drawing seem to be overridden by the dynamics of the animation. Typically, Thomas and Johnston (1984), with their canonical impact on the articulation of animation, fall short of explicitly articulating the topic of drawing. Among the twelve principles of animation, established by the Walt Disney studio in the 1930s, there is only one, *solid drawing*, that more directly addresses the drawing of the figures. Yet the authors make clear that the individual artistic skills of the animators, who provide the underlying visual form for the animation, are extremely important.

Moving further to computer generated three-dimensional renderings there is actually no line at all, only the boundary between adjacent fields or volumes. Here visual style can be discussed in terms of form, colour, texture, lighting, surface, shading, and motion, conveying a visual impression. Furthermore, experience and theories used in film and theatre seem relevant. Notably, in an attempt to extend the twelve traditional principles of animation (see above) with five additional principles applied to 3D animation, two of these five new principles are *Visual styling* and *Cinematography*. That is, both focus on visual effects (Kerlow, 2003).

An overall impression from the different contributions to the Gama Network web site *Gamasutra* for computer game developers (Gamasutra, 2004) is a frequent highlighting of the visual appearance of the characters. However, the underlying design process is primarily discussed in terms of animation principles, background stories and references such as anima style, fantasy style, Doom style, realistic style, etc., rather than as an articulation of the long process of iterative sketching and development, consisting of numerous design decisions leading up to the actual visual design.

---

**Fig. 9.** The simple inked line has tremendous visual power and almost unlimited possibilities (McCloud, 1993, p. 126). Left: © Joost Swarte; middle: © Disney Enterprises, Inc.; right: © José Munoz and Carlos Sampayo. Text by Scott McCloud (McCloud, 1993, p. 126). Reproduced with permission from HarperCollins Publishers, Inc.
6.5. Summary of the four key aspects

Taken together there is an established body of knowledge on all four aspects—degree of iconicity, face, body and costume and visual style—based on the tradition of accumulated experience in visual arts and media. In addition there are empirical studies on face and body-costume. Social psychology has a long tradition of investigating responses to different kinds of visual stereotypes.

Systematic studies of the role of such visual stereotypes in the pedagogical context of animated agents are desirable, given the possibilities to control and vary these aspects.

It is our belief that the area of new interactive pedagogical media, and specifically animated pedagogical agents, would benefit from articulating issues of visual style and including them in the research space. Furthermore, it seems that this research may be guided by a combination of the experience accumulated in the areas of animation, computer games, film and theatre together with empirical knowledge gathered in the social sciences. A caveat, however, regarding the traditions from visual arts and media is the possibility of questionable myths and misleading generalizations. An example from a related area: there seems to be a consensus among graphic designers that the lower right corner of a newspaper is an important area, something that recent eye tracking experiments demonstrate is not the case (Holmqvist and Wartenberg, submitted).

7. The broader picture and motives

It should be acknowledged that aesthetics and visual rendering are illusive aspects that are neglected not only with respect to animated agents, but in many areas of information technology design. The neglect may, however, be particularly problematic when the focus is on pedagogy and on how to motivate. Visual experiences are known to have an effect on motivation and engagingness (Laurel, 1993). In the context of animated pedagogical agents, the motivation to use and return to use a program is critical. This implies that visual experiences are too important to be treated as something to turn to when one has first solved other, more important issues.

We are certainly not denying that there are other fundamental and pressing issues that need to be solved.\(^\text{22}\)

But we maintain that visual rendering issues are pressing and fundamental, and need to be seriously approached as well. Arguments of the kind that “We are now working on the underlying engine, we first need a car that is capable of running—after that we can think about the design of the chassis” are flawed, as the parallel does not hold. Function and visual appearance are more intimately connected with one another in the case of educational programs than in the case of cars. A car has the undisputed basic function of transporting a user from one place to another. Even if a car does not visually attract or appeal to a user—or to users at all—it still, as long as it runs, performs its most fundamental task. But an educational program based on animated agents that, because of poor visualization, does not at all appeal to a user—is not motivating, has little impact—might not be used at all, even if the underlying intelligent tutoring system is well functioning. Furthermore, within the area of computer games one is repeatedly advised not to start out with any programming before the visual aspects and the story are thoroughly worked out.\(^\text{23}\)

The conceptualization of visual aspects as make-up which may be added on the surface at the end is inadequate.

Without seriously addressing aesthetic and visual aspects in research on animated pedagogical agents, the gap between academic human-computer interaction in general and experience oriented issues risks to be reproduced in this area of animated pedagogical agents. For decades, academic human-computer interaction has focused rather narrowly on usability measured in terms of efficiency and effectiveness, (e.g. Norman, 1990; Shneiderman, 1992; Nielsen, 1994; Preece et al., 1994), and the research and development methods have typically not been broad enough to deal with social and psychological factors, and even less with aesthetic and visual aspects. This is understandable, as efficiency and effectiveness are more easily approached by methods of quantitative measurement. Recently, however, some of the central researchers in the area have begun to address aesthetics as an aspect of user experience (Nielsen, 2002; Norman, 2002; Preece et al., 2002). The lesson learned in the human-computer interaction domain is that with a narrow concept of usability, it is hard to gain credibility outside of pure academia. As Cloninger (2000) puts it, this limited approach has, among others, resulted in a sometimes affected schism between academics advocating usability and graphic designers concerned with aesthetics within the multimedia and web domains. The root of the problem might be that the design and marketing of information technology artefacts deals with a wide range of product qualities, including aspects of technical features, usability and pleasure\(^\text{24}\)—and without

\(^{22}\)Not the least, issues on the intelligent tutoring system level, concerning the content of the support and the competence level of agents.

\(^{23}\)See for example the Gama Network web site Gamasutra (Gamasutra, 2004) for game developers.

\(^{24}\)In his proposal for the new human factors, Jordan (2000) describes a hierarchy of consumer needs, with functionality at the base, usability in the middle, and pleasure at the top. Pleasure is here defined as “[…] the emotional, hedonic and practical benefits associated with products” (Jordan, 1999).
methods to systematically handle the whole package, the marketing of products tends to focus on selling aspects of technical features and aesthetics, neglecting or twisting the knowledge gained in academic research, whether it is functionality, usability or something else.

With respect to animated pedagogical agent productions to be offered to schools and enterprises it is likewise important with more elaborated knowledge about the whole package, the alternative being an emphasis on selling stereotypic aesthetics and spectacular technical features.

What we are emphasizing is that it is high time to look into visual and aesthetic aspects, not waiting for the fundamentals behind to be worked out with respect to functionality, efficiency and effectiveness—and thereafter start to consider visualization issues. Recent research (Tractinsky et al., 2000) even suggests that the visual appeal of a product can influence perceived usability. There are cases where aesthetics seems to play a major role when we first form an opinion about a program, and this first opinion or judgment has an impact on the perceived quality of our subsequent interaction with the program (Lindegaard and Dudek, 2003). Furthermore, it is worth noticing when even one of the most influential names in the area of human-computer interaction, Don Norman, encourages a more aesthetic approach, by saying that “attractive things work better” (Norman, 2002).

Today much of the work is pursued in an engineering spirit, with work on implementing particular details of agents, with little regard for the goals of animated agents as a whole, that is, the goals of the design process. By asking what kinds of animated pedagogical agents we actually want to achieve, aspects of visual experiences will also be placed on the agenda. Look is too important to either be left to just happen, or to be handed over to a graphic designer outside the design team. We maintain that it is not sufficient that someone—even if it is someone with extensive know-how—fixes the visual form, without being seriously involved in the research and development process. This role ought to be integrated into the process which, in turn, may require a broadening of the design competencies in the research and development teams. There is extensive know-how among cartoonists, animators and media designers, for example, on how to engage and capture people by means of visual rendering. Unfortunately, there is rather limited formal research on qualities of look and visual rendering. A challenge for the domain of animated agents is to develop methods to handle or at least articulate the vast knowledge within such professions, so that good designs can be explained, reproduced and communicated to others.

In sum, we propose that it would be a substantial gain if scientific research on animated pedagogical agents started to address visual and aesthetic aspects seriously. The main benefit would, in the end, be to the users of pedagogical programs.

Acknowledgements

This research was supported in part by a grant from the Swedish Research Council (Vetenskapsrådet).

The authors would like to thank the two anonymous reviewers for highly valuable comments on a previous version of this paper.

References


Holmqvist, K., Wartenberg, C. The role of local design factors for newspaper reading behaviour—an eye-tracking perspective (submitted).


