Virtual Pedagogical Agents
Beyond the Constraints of the Computational Approach

Magnus Haake

Licentiate Thesis
Div. of Ergonomics and Aerosol Technology
Dept. of Design Sciences
Lund University
Sweden
To Agneta – my everything …

… and to Sebastian and Igis – who, with their childish egoism, never gave me a real chance to confuse the reality of life with the virtuality of academia.
Virtual Pedagogical Agents
Beyond the Constraints of the Computational Approach

Magnus Haake

Licentiate Thesis
Ergonomics and Aerosol Technology (EAT)
Department of Design Sciences, LTH
Lund University, Sweden
Virtual Pedagogical Agents
– Beyond the Constraints of the Computational Approach

Copyright © 2006 by Magnus Haake
Typefaces: Frutiger and AGaramond

Ergonomics and Aerosol Technology (EAT)
Dept. of Design Sciences, LTH
Lund University
Sweden
http://www.design.lth.se/

ISSN 1650-9773
Publication 24
ISRN LUTMDN/TMAT-2028-SE
EAT 2006

Printed in Sweden by MediaTryck, Lund, Sweden
# Table of Contents

Abstract ............................................................ V  
Sammanfattning ....................................................... VII  
List of Included Papers ............................................... IX  
Other Work by the Respondent .................................. XII  

**CHAPTER 1: VIRTUAL EMBODIED PEDAGOGICAL AGENTS** .................................... 1  
Virtual embodied pedagogical agents ................................. 1  

**CHAPTER 2: VIRTUAL PEDAGOGICAL AGENTS – A SHORT HISTORICAL RÉSUMÉ** .......... 5  
Knowledge systems and artificial intelligence ........................ 5  
User modelling ......................................................... 7  
Embodiment ................................................................ 7  
Computer based pedagogy ............................................. 7  
Present state ............................................................. 8  
References ................................................................ 9  

**CHAPTER 3: A RESEARCH OVERVIEW OF VIRTUAL PEDAGOGICAL AGENTS** .......... 10  
A new social actor ..................................................... 10  
The potential to engage and motivate ................................ 11  
Literature on visual appearance in relation to virtual agents .... 13  
Summary and discussion ............................................... 14  

**CHAPTER 4: RESEARCH QUESTIONS** ................................................................. 17  
The need for visual design guidelines ................................. 17  
The research questions .................................................. 18  

**CHAPTER 5: VISUAL DESIGN AND VIRTUAL PEDAGOGICAL AGENTS** .................. 20  
Do visual aspects matter? .............................................. 20  
The importance of visual aspects ..................................... 20  
Explanations for the neglect of visual design issues ............... 22  
Conclusion and discussion: Research question # 1 ............... 23  

**CHAPTER 6: VISUAL ASPECTS AND PEDAGOGICAL ISSUES** ............................... 25  
Iconization and subjective involvement ............................ 25  
The two studies ......................................................... 26  
Results and conclusion: Research question # 2 ................. 28  

**CHAPTER 7: VISUAL STEREOTYPES AND VIRTUAL PEDAGOGICAL AGENTS** .......... 29  
Visual aspects .......................................................... 29  
Visual stereotypes .................................................... 30  
Cognitive implications of visual stereotypes ....................... 30  
Visual stereotypes and virtual pedagogical agents ............... 31  
Summary: Research question # 3 .................................. 32
CHAPTER 8: SUMMARY OF THE THREE INCLUDED PAPERS ........................................34
    Summary ........................................................................................................34
CHAPTER 9: DISCUSSION & FUTURE POSSIBILITIES ........................................37
    The visual design space of virtual agents .......................................................37
    A criticism of user evaluations ....................................................................40
    Virtual ‘assistants’ in the health domain .......................................................40
    Virtual agents as a research tool ..................................................................41
    Future research agenda ..............................................................................42
REFERENCES ......................................................................................................45
    Virtual pedagogical agents, VPAs (Page 3) .................................................45
    Computer assisted learning systems, CALs ...............................................46
    Articles ........................................................................................................46
Acknowledgements

I have a position that is becoming rare in Sweden. As a faculty employed doctoral student with no specific project funding, I find myself with the opportunity to pursue research out of pure curiosity. For this I am very much obliged to the otherwise most peculiar Swedish academic system, where most people seem to carry out their research not thanks to, but in spite of, the system.

Of course I also want to give my regards to the people around me at the department of Design Sciences in Lund, as well as to the people at LUdS (Cognitive Science Lund University). In particular I want to thank our marvellous secretary Karin Öhrvik for helping us all with all our stupid problems, and Mattias Wallergård for his interest in my virtual creatures, his tips concerning film and computer games, and his patience when I complain over all sorts of things.

As to this very thesis I owe a lot to my wife Agneta, who has been my research colleague for years. By now, our academic engagements are totally intermixed – and in some sense we could be regarded as a strange symbiotic academic specimen.

I must not forget my children, Sebastian and Igis. They have had their share of strange conversations around the dinner table. Especially, I remember Sebastian asking, ‘Why has everything something with design in it?’

Magnus Haake
Lund, 24/10 2006
Abstract

Virtual pedagogical agents (VPAs), i.e. computer generated characters in pedagogical roles, populate the digital society in increasing numbers. They are found in educational programs from preschool to university. In a broader educational context, they function as virtual medical counsellors, physical exercise coaches and guides on city web pages – and they also appear in edutainment and infotainment settings.

The main reason for adding virtual (pedagogical) agents to educational software is their potential to motivate and engage. In order to develop and exploit this potential, there are several issues that need to be resolved. This thesis (based on three papers) begins with an overview of research on motivation and engagement in VPAs. The outcome of this overview is that visual aspects of VPAs in much are neglected.

The thesis continues with an argumentation on the importance of visual aspects in VPAs, followed by a presentation of two empirical studies. The two studies examine possible correlations between visual and social/communicative characteristics in VPAs and social/communicative styles in users. Taken together, there is strong support for the claim that users’ visual/aesthetic experience of VPAs is too important with respect to the goals to motivate and engage, to be treated as a secondary issue.

Thereafter follows a survey on visual properties of VPAs in relation to visual stereotypes and novel possibilities as well as risks are discussed.

The conclusions of this thesis are: (i) that it is highly probable that visual aspects in VPAs can be related to pedagogical outcomes; (ii) that there may be significant relations between visual and social characteristics of VPAs and social/communicative styles of learners; and (iii) that VPAs may reproduce stereotypes from everyday real life human-human interaction, as well as from traditional visual media – something that, in combination with the increased interactive potential of VPAs, may even strengthen certain detrimental effects of idealized stereotypical instances.

The research agenda proposes a further exploration of visually related aspects of VPAs – not least in order to handle the outcomes of user evaluations that otherwise will continue to be blurred by the impact of uncontrolled visually related variables. Furthermore, the research agenda argues that there is a problem with user evaluations based upon the concept of ‘the User’, i.e. the averaged standardized user. The reason for this is
that users may diverge into categories with almost diametrically opposite qualities as to, for example, personality, communicative style and learning style. In order to reveal important correlations between agent characteristics and user characteristics, it is therefore necessary to identify such groups of users.

The research agenda also points at the need for ethical and legal research regarding VPAs, as there now is a golden opportunity to attempt to address such questions in advance.

And finally – virtual agents may have a not yet thought of potential as research tools.
Virtuella pedagogiska agenter, dvs. datorgenererade karaktärer i pedagogiska sammanhang, har börjat invadera den digitala världen. De dyker upp i utbildningsprogram från förskolan till universitetet. I ett bredare utbildningssammanhang kan de fungera som tränare, virtuella medicinska rådgivare eller guider på olika kommuners hemsidor – och de förekommer också inom edutainment och infotainment.

Ett huvudsakligt till att använda virtuella (pedagogiska) agenter i utbildningsprogram är deras potential att motivera och engagera. För att utveckla och utnyttja denna potential behövs vidare forskning. Denna licentiatuppsats (som baseras på tre artiklar) börjar med en översikt och genomgång av forskningen kring motivation och engagemang i samband med virtuella (pedagogiska) agenter. Genomgången visar att visuella aspekter i stort har negligerats.

Uppsatser fortsätter med en genomgång av visuella aspekter hos virtuella pedagogiska agenter. Därefter presenteras två empiriska studier som utvärderar sambandet mellan virtuella pedagogiska agenters visuella och sociala stil i relation till användares sociala stil. Sammantaget finns ett starkt stöd för påståendet att användares visuella/estetiska upplevelse av virtuella pedagogiska agenter är alltför viktigt för att negligeras.

Den sista delen av uppsatsen diskuterar virtuella pedagogiska agenter med avseende på visuella stereotyper och visar på framtida möjligheter och risker.

Slutledningarna i uppsatsen är: (i) att det är troligt att visuella aspekter hos virtuella pedagogiska agenter kan kopplas till pedagogiska effekter; (ii) att det finns signifikanta samband mellan visuell och social stil hos en agent och social stil hos användare; och (iii) att virtuella pedagogiska agenter reproducerar stereotyper (fördomar) från den dagliga interaktionen mellan människor, såväl som stereotyper från traditionell media – och att de utökade interaktiva möjligheterna med virtuella pedagogiska agenter kan förstärka negativa effekter med avseende på idealiserade stereotyper.

Förslagen till framtida fortsatt forskning lyfter fram behovet att utforska visuella aspekter hos virtuella pedagogiska agenter – inte minst med tanke på användarutvärderingar, vilka annars riskerar att även fortsättningsvis försvagas av okontrollerade effekter orsakade av visuella egenskaper. Där till är det nödvändigt att identifiera relevanta användarkategorier för att...
kunna upptäcka viktiga samband mellan egenskaper hos virtuella pedagogiska agenter respektive användare.

Forskningsförslagen argumenterar också för att en framväxt av virtuella pedagogiska agenter kommer att medföra etiska och juridiska komplikationer.

Sist – men inte minst – kan virtuella pedagogiska agenter skomma att få en spridd användning som forskningsverktyg och forskningsmetod.
List of Included Papers

_This thesis is based upon the following three papers, which will be referred to in the text by their Roman numerals (Paper I, II & III)._  
Both authors have been involved in the work of these three papers (Paper I, II & III) – Agneta Gulz being more experienced in computer mediated pedagogy and Magnus Haake being more experienced in visualization.  
As to the years for the included papers, Paper II is dated to ‘2005’ and Paper I to ‘2006’. This is due to the often extended process of publishing in a Journal (Paper I) in comparison to a Proceeding (Paper II). The order of the papers, however, reflects the academic process.

I. **Design of Animated Pedagogical Agents – a Look at Their Look**  
(NB: In Paper I, the term ‘animated’ is frequently used. In later works – and in this thesis – ‘virtual’ is consequently used in favour of ‘animated’.)

**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. 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.

II. Social and Visual Style in Virtual Pedagogical Agents


(NB: In the thesis, the terms ‘naturalism’ and ‘realism’ are considered to be equivalent, as are ‘stylization’ and ‘iconization’.)

Abstract: The paper addresses aspects of virtual pedagogical agents’ visual style (realism – iconization) in relation to their social style (task oriented – relation oriented). Two studies are presented that investigate which visual and social styles users prefer and how they articulate their preferences. The first study involved 42 university students; the second study involved 90 elementary school children. Special emphasis was put upon two hypotheses, grounded in cognitive theory: (i) iconized visualization may be better suited for representing a relation oriented, subjective agent – and therefore preferred by users who prefer a relation oriented agent; (ii) realistic visualization may be better suited for representing a task oriented, objective agent – and therefore preferred by users who prefer a task oriented agent.

The results of the two studies provide some support to these hypotheses. Cognitive theories are exploited to interpret the results, and possible design considerations are discussed.

III. Aesthetic Stereotypes and Virtual Pedagogical Agents


(NB: This paper was submitted to the Journal of Educational Technology & Society in June, 2006.)

Abstract: The paper deals with the use of visual stereotypes in virtual pedagogical agents and its potential impact in digital electronic environments. After an analysis of the concept of visual stereotypes, affordances as well as drawbacks of their use in the context of traditional media are analyzed. Next, the paper explores whether virtual
pedagogical characters introduce anything novel with respect to the use of visual stereotypes – as compared both to real life interaction between humans and to the use of visual stereotypes in traditional non-interactive media such as magazines, film, TV and video. The answer is that novel affordances, as well as novel drawbacks, indeed are being introduced with the use of visual stereotypes in virtual characters. Finally, we suggest some future directions of research, and from an educational and societal perspective we argue for the imperativeness of such research.
Other Work by the Respondent


INTRODUCTION &
RESEARCH OVERVIEW
CHAPTER 1: VIRTUAL EMBODIED PEDAGOGICAL AGENTS

You find yourself in a group of people you have not met with before – and engage in the usual small talk. Inevitably someone asks you what you work with. You hesitate for a second, and then reply, ‘Virtual embodied pedagogical agents …’

The response is most often confusion, and therefore you start to explain, ‘I work with computerized animated characters that act as virtual teachers or assistants.’

The response may now change from uncertain confusion to something like fear. You may even face indignant or agitated accusations, as if you were involved in something of utmost immorality. As the situation grows uncomfortable, you quickly add, ‘Like the paper clip in Microsoft Office.’

… relief!

Figure 1. The (in)famous Microsoft Office Assistant ‘Clippit’. (© Microsoft Corporation)

Virtual embodied pedagogical agents

In order to explain what a virtual embodied pedagogical agent is, the assistant in Microsoft Office is a good start. It is a tangible example that most people in the computerized parts of the world have some experience with. The fact that the assistant so clearly lacks any kind of actual intelligence and autonomy also wipes away the uneasiness that may derive from the diffuse threat of artificial life – a frequently recurring theme in science fiction with murderous robots taking over the world.

But even if the Microsoft Office Assistant constitutes no threat, people are affected. Many hate it, while others appreciate it. When it first turned up in the late 1990s, the Internet was flooded with desperate cries, asking
for how to get rid of the stupid and irritating little creature that continuously insisted on showing up at the wrong moments, suggesting help with the wrong problems and refusing to disappear. (At that time it was quite tricky to get rid of the assistant.)

In any case, the virtual assistant in Microsoft Office actually has the basic characteristics that constitute a virtual embodied pedagogical agent:

1. It is built upon an interactive knowledge system.
2. It has a model of the user.
3. It is embodied, i.e. given a visual ‘gestalt’.
4. A pedagogical approach underlies the characteristics of (1) to (3).

This list of characteristics also outlines the evolution of virtual embodied pedagogical agents, which will be presented in the next chapter. But first, a closer look at the concept of virtual embodied pedagogical agents:

- **Agent** means (from a computational perspective) a computer program that is in some sense autonomous. In the context of virtual pedagogical agents, this implies some kind of simulated autonomous behaviour in the interaction with the learner.

- **Pedagogical** refers to the educational framework, for which the computer based program is developed.

- **Embodied** points to the fact that the computer program is represented through a visual character. Something that differentiates these systems from other pedagogical systems where, for example, the interaction takes place only through textual input and output. Virtual agents, however, are not embodied in the more narrow sense of having a *physical* body (c.f. embodied cognition). In that sense a robot, but not a virtual agent, is embodied.

- **Virtual** emphasizes the non-physical, digital computer generated context (c.f. virtual reality).

The notion of ‘embodiment’ is, however, often left out even if implied, resulting in the common expression *virtual pedagogical agents (VPAs)*. Another frequently used term is *embodied conversational agents (ECAs)*, representing virtual embodied agents that can engage users in multimodal conversation including both verbal and non-verbal behaviour, where the conversations in question may or may not take place in pedagogical contexts.
Other additional terms for designating these computer generated characters – with or without a pedagogical aim – are: *animated agents*, *virtual characters*, *synthetic characters* and *chatbots*.

**Examples of virtual pedagogical agents**

There are a steadily increasing number of educational systems and prototypes that use different kinds of virtual pedagogical agents. The agents appear as virtual teachers, instructors, coaches, learning companions, presenters, guides, and also work together in teams (Figure 2a & 2b).

Figure 2a. VPAs as teachers, instructors or coaches

![AutoTutor](image1)  
*AutoTutor*  
CSL, Univ. of Memphis  

![Herman the Bug](image2)  
*Herman the Bug*  
IntelliMedia  

![Laura](image3)  
*Laura*  
MIT Media Lab  

![Steve](image4)  
*Steve*  
CARTE  

![Sam](image5)  
*Sam*  
GNL, MIT Media Lab  

![FearNot](image6)  
*FearNot*  
eCIRCUS /VICTEC  

Figure 2b. VPAs as learning companions

![Olga](image7)  
*Olga*  
KTH, SU, SICS, Nordvis  

![Rea](image8)  
*Rea*  
GNL, MIT Media Lab  

![Anna](image9)  
*Anna*  
IKEA United Kingdom  

Figure 2c. VPAs as presenters or guides (assistants)
The last examples (Figure 2c) illustrate virtual agents that function as ‘presenters or guides’. These are not properly educational, in the sense that they are implemented in systems that are not clearly educational. At the same time, these systems incorporate teaching elements in that they, for example guide the user through sets of information.
CHAPTER 2: VIRTUAL PEDAGOGICAL AGENTS – A SHORT HISTORICAL RÉSUMÉ

This brief historical overview illustrates a personal perspective on virtual pedagogical agents (VPAs). One can take other perspectives that would tell a different story, thus complementing and broadening the bigger picture. The overview is also a strongly oversimplified version of the academic progress and evolution in this area over four decades. However, I consider this résumé accurate enough for a basic understanding of virtual pedagogical agents.

Knowledge systems and artificial intelligence

With the growing interest and popularity for artificial intelligence in the 1960s, an early vision of machine-assisted learning emerged. The idea was to use artificial intelligence to provide students with a personal and individualized tutor – an intelligent tutoring system (ITS). Such a system should be built upon three components: domain knowledge (knowledge of the specific subject), student model (a way to ‘understand’ and interact with the student), and teaching knowledge (pedagogy). In this way, the system would be able to survey a student’s actions and progress, provide feedback and give contextual advice and support for problem solving.

Figure 3. A marvellous illustration found in a psychology book from the 1960s. (Psychology Today, 1967; Drawing by Don Wright.)
During the 1960s and 1970s, the first computer assisted learning systems (CALs) were developed, such as *PLATO IV* (1972) [1] and *TICCIT* (1975) [2]. These systems were more of ‘knowledge systems’, interacting with the user through text based input and output (keyboard and screen). Their so called intelligence was based upon domain specific data, systemized in databases with simple rules for interaction with the student. They could prompt the user through a static tutoring plan and answer user questions by means of a huge database (as long as the user typed the right question with the right syntax). It was all on the terms of the system (though developed by humans) and needless to say – not a success.

This initial failure was in many ways related to the misconception of artificial intelligence and human cognition. During the 1950s and especially during the AI-hype in the 1960s and 70s, there was a widespread optimism for the possibility to describe the world in general rule based terms. It was rational, logical thinking that was held to separate humans from all other biological forms of life, and consequently artificial intelligence should mimic the rational, logical process. Actually, this was relatively easily accomplished. However, these ‘intelligent systems’ did not turn out a success, for instance in decision making. Today it is easy to joke about the rational paradigm, but though it gradually declined as a general scientific approach, the mismatches between goals and accomplishments in this paradigm have provided much useful knowledge.

**A new cognitive approach**

The rational paradigm could not provide any substantial explanatory models for human cognition. Humans may be separated from animals by their cognitive capabilities, but these are not strictly rational and logical. It is even questionable whether humans can process a simple logical problem in a distinct segment of the brain without affecting other areas of the brain to different degrees. For instance, a person may answer that ‘2+3 equals 4’ just because he or she feels like doing so – which goes beyond any rational model.

This irrationality has not been easy to capture – at least not for the positivistic and rationalistic AI approach. As this approach cooled down, a more complex picture emerged with a complex, gradually evolved brain situated in a continuous and highly reciprocal system of brain and body as well as a surrounding world. This view of cognition as situated and embodied is far from Descartes’ separation of body and soul. Actually
a human brain not ‘stimulated’ by external signals would languish away – after having probably passed through a psychotic stage).

**User modelling**

A lesson learnt from the progress of the cognitive sciences is that a strictly rule based and domain specific tutoring system will most certainly run into conflicts with irrational and context dependent humans. Consequently, an intelligent tutoring system needs a strategy to handle such phenomena as emotions, affect, motivation and engagement, i.e. there is a need for computational *models* of the human user.

**Embodiment**

Having (at least hypothetically) settled for an intelligent tutoring system (ITS) that can adapt to human ‘irrationalities’, another question appears: How should interaction between the human and the ITS take place? In line with the reasoning above, humans are relatively bad or unreliable when communicating through the rule based and constrained textual languages of computers compared to their skill at communicating in spoken language, including eye gaze, facial gestures and body language. In other words, there is a need to support human communication skills. Adding a body as a visual graphical element to the ITS is one way to try to adapt to multimodal human communication.

**Computer based pedagogy**

The evolution of virtual pedagogical agents is intermixed with a parallel development in pedagogy, in part due to the relation that both fields have to the cognitive sciences.

In contrast to a ‘traditional’ view, more based on the individual plodding through facts and the training of rational thinking, pedagogy of today conceives of learning as basically a social activity that involves interaction and communication with mentors, teachers and fellow students in the development of one’s own understanding and learning skills.

Given the ambition in many parts of the world to educate a mayor part of the population all the way up to university levels, there is a need for new strategies if this vision or concept of learning is to be implemented.
One such possible approach is the use of information technology in the form of discussion groups, on-line debates, chats, forums, etc. Such systems provide the infrastructure for different social exercises, where the social contribution mainly is provided by the participants themselves.

But as an expansion of the education system also results in a shortage of teachers, and especially good teachers, another possible approach is the use of virtual pedagogical agents with social and emotional skills – which, once created, are inexhaustible and possible to clone in infinite numbers.

**Present state**

As described, virtual pedagogical agents has a heritage in the positivistic vision of artificial intelligence in the 1960s. As a consequence of the decline of the old AI paradigm and the revision of human cognition, the further development of intelligent tutoring systems had to incorporate knowledge from disciplines such as cognitive science, computer linguistics, psychology, social science, pedagogy and graphic design.

Yet, despite the incorporation of disciplines from the humanities and the social sciences, the discipline of virtual pedagogical agents is still basically computer oriented. This is only natural, as a virtual pedagogical agent together with an educational system is a digital interactive artefact. The drawback is that perspectives from other disciplines that are not *directly computational* are more or less suppressed.

The present status of virtual pedagogical agents is puzzling. Different kinds of virtual (pedagogical) agents appear on the web in the form of ‘chatbots’, e.g. the virtual assistant Anna, helping visitors at IKEA’s web sites all over the world (Figure 2c). Furthermore, there are the virtual pedagogical assistants long since established in CD-ROMs for children and training programs for companies. Yet, if one takes a closer look at these figures, they do not incorporate much autonomy and/or intelligence and thus are not true *agents* according to the definition. They are also strictly staged with sparse animation, and only deliver a constrained set of pre-recorded comments.

When it comes to the more advanced virtual pedagogical agents found in research contexts, there are various agents specialized to address different research questions. Many of these indeed show promising results, but they often are very specialized and the generalization of the results lies in the future. There are a few instances where fully functional agents are used
to deal with real world problems (e.g. *Laura* [4], *FearNot* [5], and a set of virtual speech therapists [3]). Even though these, as well, are primarily research prototypes, the results are promising and commercialization is within reach.

In conclusion, today it is possible to develop a virtual pedagogical agents with simple and rudimentary models of human behaviour.

**References**

Much of what is reported in this first chapter is a heavily compressed mixture of readings, comments and arguments I have come across during the more than fifteen years I have spent in the areas of HCI/Interaction Design and Cognitive Sciences. My perspective is coloured by the assumption that, in spite of all the scans, flow charts and computational models of the human brain, we are far from an actual understanding of the brain with its evolutionary heritage, tightly connected to the body and continuously interacting with the surrounding world. Today we can study different cognitive phenomena, but the data only tell us *that* there is an increased neurological activity and *where* most of this activity take place – not *why* and *how*.

As for references, I suggest the following books and articles for those who want to read more:

CHAPTER 3: A RESEARCH OVERVIEW OF VIRTUAL PEDAGOGICAL AGENTS

Paper I of this thesis (Design of Animated Pedagogical Agents – a Look at Their Look) has an elaborated overview of virtual pedagogical agents and the research area. What follows is a condensed and updated version.

A new social actor

As mentioned in the previous chapter, virtual pedagogical agents can be traced back to the intelligent tutoring systems of the 1970s. The basic idea then, was to provide students with a personal and individualized tutoring system, able to survey a student’s actions and progress, as well as provide feedback and advice.

This vision is still alive today, though the approach is more humble, and not least the modelling of learners’ behaviour is today recognized as being problematic.

The now classical intelligent tutoring system was invisible and abstract, communicating by textual input and output. Thus, the addition of a virtual conversational agent to the interface of the tutoring system opens up for a more interesting and effective approach. The virtual agent gives the system ‘a face’, which makes it easier for the student to interact with it, as if it had socio-emotive abilities and even a personality, communicated by means of gestures, facial expressions and clothing. In this way, the user can construct a personal and emotional relationship to an agent-system, thus creating a social context which will be beneficial for learning according to the theory of learning as a social activity.

This is not just speculative science fiction. In 1996, Byron Reeves and Clifford Nass in their seminal book, The Media Equation: How People Treat Computers, Televisions and New Media Like Real People and Places [54], convincingly showed that people in many aspects behave towards computer systems as they do towards real people. As an example, people are affected by a computer that flatters; and likewise, people are more moderate in their criticism if an evaluation of a computer system is performed at the same computer (as the tested system) compared to filling in a paper questionnaire in another room.
A plausible explanation is that humans by evolution are skilled in socially communicating with other humans, using communicative strategies that often are so automated that they do not (and need not) think about them. From an academic point of view, these ways of communicating may seem fuzzy and elusive, but from a cognitive perspective they are extremely efficient and economical, particularly compared to the precise and structured interaction with computer systems.

If the computer system then offers an even vague hint of human characteristics and behaviour, people easily and unconsciously fall back on these human communication strategies. In other words, people easily establish emotional and personal relations to computer systems – and if a visible character is added, no matter how simple and rudimentary the models of emotions and behaviour may be, people will respond to it.

The potential to engage and motivate

A central motive for the development of virtual pedagogical agents is to enhance learning in students and this puts the light on the strong, albeit complex, relationship between motivation, engagement and learning.

Engagement

One of the main motives for the introduction of a virtual agent in pedagogical contexts is to incorporate an experience of engagement in the learning situation. An increase in engagement may encourage a student to interact with a pedagogical system more frequently and increase the time spent with a system, which may result in improved learning outcomes.

Engagement is, furthermore, the most well established effect of virtual agents [7] [28] [38] [47] [55] [63] [67] [70]. A way to evaluate ‘engagement’ is to study the relation between attitude measures (e.g. the extent to which participants enjoy the experience and/or feel involved in the interaction) and behavioural measures (e.g. how long do the participants stay on in a learning environment, how much activity does s/he exhibit, how willing is s/he to use the program again). If a program is found to be engaging, it is likely that users will be more active and stay on longer. However, engagingness is not to be equated with entertainment, as one can be involved without feeling amused.

Several studies have also demonstrated that users react differently to virtual agents based on their own personality and other dispositional traits.
– the engagingness effect, in particular, varies. For instance, Reeves and Nass [54] and Nass et al. [50] showed that users prefer agents that match their own personality on the introversion – extroversion dimension.

Motivation
The research issues regarding motivational effects have developed since the early evaluations of general motivational effects in favour of an increasing concern for which people, in which conditions and for what kinds of domains that a certain effect can be shown [38] [55] [67]. As an example, van Mulken et al. [67] showed that the presence of a virtual agent was more favourable in a pedagogical technical context, than in a introduction of employees in a company. In the latter case, the photos of the employees probably evoked the sought after social effect, and the agent did not add to this effect.

A way to approach motivation is to look at the three aspects of how motivation may affect learning suggested by Schank and Neaman [58]. The first aspect is participation, e.g. whether a student participates at all in a learning activity. The second is indexing, which can be understood as the organization of memories. The third is attention, which affects what is remembered. (The attention aspect is complex, as too much attention may distract from learning.)

Design elements of virtual agents
In pursuing research on effects of motivation and engagingness, it is necessary explore the elements of virtual pedagogical agents that are likely to effect user experiences. Several such elements have indeed been studied:
– Movement characteristics, in particular gestures and hand movements [24] [39] [40] [41].
– Facial expressions [40] [41] [53].
– Voice characteristics [8] [49] [51].
– Dialogue and conversational characteristics [23] [44].
– Emotional expression via voice, gestures, facial expressions [9] [10] [14] [40] [41].
– Personality realized via voice, gestures, facial expression, verbal communication [8] [10] [25] [33] [64].
This is a long list, and numerous references may be added. Nevertheless,
a fundamental aspect of virtual agents is missing, namely, the underlying visual gestalt. This neglected aspect is the focus of this thesis, and a presentation of the sparse agent-related research on visual form follows.

**Literature on visual appearance in relation to virtual agents**

Compared to the amount of research on the other design aspects (movements, facial expressions, voice, conversation, emotions and personality) little that has been carried out on the visual aspects of virtual pedagogical agents. With respect to the central pedagogical potential of virtual pedagogical agents to be engaging and motivating – to increase involvement and contribute to the impact of learning activities – this research gap may be unfortunate.

**Personality & visual appearance**

In research on virtual pedagogical agents, the personalities of the agents surfaces as a central issue [33], and one important rationale for personality is its significance for motivation and engagingness. But whereas there is a bulk of research on the relation between personality and verbal and non-verbal communication, personality and visual appearance is hardly addressed. André et al. [8], however, made an remark in that both the visual appearance and voice of a character are important cues to its personality and interest profile. As a consequence, the possibility to reuse the visual appearance and voice of characters for different roles is limited.

**Facial expressions & visual appearance**

Another key issue in research on virtual pedagogical agents is facial expressions. Typically, the research has focused on computational oriented problems like ‘real time’ generation of appropriate and believable expressions. Thus far, there are no reports of systematic studies on whether a specific algorithm will produce consistent or diverse results depending on controlled visual variations of the underlying face.

**Gestures & visual appearance**

Whereas qualities such as gestures and movements are extensively addressed, the underlying face, body and clothing – on which gestures and movements are superimposed – is little articulated. This is a remarkable contrast to the design process in animated cartoons and games, where
an extended period of time may be dedicated to the iterative process of sketching and evaluations to decide upon the right face, body, costume and specific attributes [65].

Summary and discussion

Issues of visual aspects are much neglected in virtual agent research. When it comes to systematic studies of visual aspects in pedagogical agents, Amy Baylor and her group at the Centre for Research of Innovative Technologies for Learning (RITL), Florida State University, are almost the exception that proves the rule [11] [12] [13].

At the same time, there is much knowledge in the game industry, but it is more of an accumulated and implicit knowledge that is not easily adaptable to the academic process. For a long time, the frequently cited bible in this domain has been The Illusion of Life: Disney Animation by Disney animators Frank Thomas and Ollie Johnston from 1984 [65]. Though an impressive work, the focus of the book is on animation, and one has to read between the lines (and illustrations) to find anything about the basic underlying visual appearance of the characters. Only as recent as 2006, Katherine Isbister published her book, Better Game Characters by Design: A Psychological Approach [32], that more explicitly examines the basic design of the characters themselves.

Turning to the areas of film and theatre, the craft of staging and dressing has a longer, more elaborated tradition. They have also established themselves as academic disciplines with their own discourse. Most likely, it would be beneficial to incorporate knowledge from these traditions in the development of virtual pedagogical agents.

All in all, there is considerable knowledge, mostly implicit but also to some extent explicit, which concerns visual aspects. A reminder though: some of the general recognized rules related to character design may be false. There seems to be a consensus among graphic designers, for example, 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 [31].
- Research -
A Résumé of the Included Papers
CHAPTER 4: RESEARCH QUESTIONS

The preceding chapters have briefly covered the concept of virtual (embodied) pedagogical agents and their rather short history. Even if the chronicle of virtual pedagogical agents still has to be written, these computer generated figures engaged in pedagogical roles are already entering the digital society in increasing numbers. They are found in educational programs, from preschool to university. They are also found in broader educational contexts as virtual medical counsellors, physical exercise coaches and guides on city web pages — and (not least) they appear in edutainment and infotainment settings.

The need for visual design guidelines

In the preceding Section, the research survey made it obvious that empirical research on visual design aspects are rare. Consequently, there are no adequate design guidelines as to visual matters.

An intuitive way to go in order to generate such guidelines may be to look at the domains of game industry, film and performing arts, that are already in possession of substantial accumulated experience and knowledge regarding character design. However, much of this is tacit knowledge — hard or even impossible to capture in words and general rules. Furthermore, it is most uncertain how transferable and applicable this experience and knowledge is in a pedagogical context. In a computer game, for example, one wants the user to feel good in order to generate a positive attitude to the game [32]. This means that characters more or less inevitably must be designed to be stereotypical ‘attractive’ — this often also goes for the bad guys. If then the game itself is ‘entertaining’ — everything is fine. But if one is to present a virtual pedagogical agent in a more or less mandatory teaching context to a heterogeneous group of ambivalent, contradictory and hormone distended teenagers, how should this be done?

What knowledge is there to be relied upon when facing the need to make adequate design decisions in the visual design of a new virtual pedagogical agent? Should one rely on guidelines for the ‘feel good’ approach mentioned above? Or exploit common stereotypical instances, such as the serious and rather thin male character wearing spectacles who teaches physics or hard wired engineering subjects? Or should one design an agent
with bee-sting lips, a micro-thin waist and voluminous, pneumatic breasts to obtain the interest of teenaged boys?

When it comes to educational software, several problems may arise with a design strategy originated from an entertainment based approach. As shown by Amy Baylor and her group at RITL [11], students may feel more comfortable and show a more positive attitude with a virtual teacher that is similar to themselves in terms of gender, ethnicity and/or social position. But at the same time, students may perform better if the virtual teacher belongs to a group having a lower social position. The underlying reason is that students from a ‘higher’ social position work harder as they experience a threat against the social hierarchy they benefit from.

If the goal then is to maximize a positive attitude, one might want to rely on identification, thereby sanctioning established prejudices. If the goal, on the other hand, is to maximize the learning outcome, one may consider using ingrained and inferior social stereotypes as to gender and ethnicity. Both goals are desirable, while the strategies are morally/ethically, as well as socially and politically questionable.

Another issue is that more extensive ventures in agent-based educational software for schools would be a part of a governmental agenda, thus reflecting and addressing ideological standpoints of the government and society. To what degree can agent-based educational media be used (and misused) to reflect these standpoints? Could a hypothetical governmental design guideline be used, for instance, to come up with a desexualized female teaching auto repair and machine maintenance? This is in many ways as troublesome as the utilitarian approach discussed in the previous paragraph. Compare this to the modernistic social engineering paradigm in Sweden during the 1940s and 50s [56]. The idea to ‘educate and enlighten’ people by designing their lives down to the tiniest detail would probably only provoke a desire to oppose or to be incorrect – just for the satisfaction of not letting others decide over one’s life.

The question remains: ‘What should visual design guidelines for virtual pedagogical agents look like?’

**The research questions**

Before trying to approach the complex and controversial question just discussed, it is necessary to go back to the basic theme of this thesis: ‘Do visual aspects of virtual characters matter, and if so – how?’
– According to the research of Byron Reeves and Clifford Nass [54], people reproduce (stereotypical) behaviour and strategies of human-human interaction into human-computer interaction.

– Amy Baylor and her group at RITL [11] have shown that people judge virtual pedagogical agents stereotypically according to their visual appearance in terms of gender and ethnicity.

– Katherine Isbister [32] argues convincingly for the importance of graphical design of virtual characters in game design.

The answer to the question is then: Visual aspects of virtual pedagogical agents do matter – and humans respond to these agents in ways similar to the ways they respond to real humans.

In view of this, it is time to once again approach the question of how to design virtual pedagogical agents. Any kind of design agenda for virtual pedagogical agents is a complex and controversial issue, as it will inevitably be influenced by social and political standpoints. From this point of view, scientific research is not in possession of the answer to these questions. However it may uncover and dismantle important aspects and connections. In this way scientific research may influence the political agenda and contribute to positive progress.

In line with this, the aim of this thesis is to explore this area by addressing the following questions:

(1) Does the visual appearance or representation of a virtual pedagogical agent matter (from a pedagogical point of view)? [Paper I]

(2) Is there some kind of relation between visual characteristics of virtual pedagogical agents and characteristics of the learner? [Paper II]

(3) If the answers to questions (1) and (2) are affirmative, what are their consequences, given the impact of visual stereotypes in the society? [Paper III]

The following three chapters (corresponding to the three papers) in turn address the three questions above.
CHAPTER 5: VISUAL DESIGN AND VIRTUAL PEDAGOGICAL AGENTS

Paper I ‘Design of Animated Pedagogical Agents – a Look at their Look’, addresses the first research question: ‘Does the visual appearance or representation of a virtual pedagogical agent matter (from a pedagogical point of view)?’ The first part of the paper consists of a research overview, outlined in Chapter 3 in the previous Section. What follows is a summary of the second part, somewhat reformulated and with updated references.

Do visual aspects matter?

During the process of reviewing the research area of virtual pedagogical agents, it soon became clear that visual or graphical issues were left out, for the most part, in favour of computational oriented approaches. In particular, there were hardly any comparative studies that systematically varied graphical dimensions.

At the same time, it was indeed hard to conclude that the visual representation of pedagogical agents should not matter given the experience and knowledge from the game industry, films, performing arts, cartoons and social psychology. Rather, with respect to the potential of virtual agents to engage and motivate in pedagogical contexts, the question was reformulated: To what extent does visual appearance matter? When, where and how does it matter?

The importance of visual aspects

It has long been established in social psychology that other people’s appearances and observable physical cues, such as clothing, profoundly affect our judgements [33] [44] [47] [60] [61]. It has also been shown that the same holds for virtual agents in many cases [30] [37] [52] [63] – and this also counts for pedagogical settings [10] [11] [12] [29].

A plausible argument for the impact of visual appearance is its role in the representation of personality. Branham [19] 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, gender, race, physical attractiveness, hair, clothing, make-up, facial shape, facial hair, and so on. Likewise, Berscheid and Walster [14] 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.’ The crucial issue is that regardless of how accurate such insights are, people do build them.

![Stereotypical drawing based on different characters in the Swedish Disney magazine Prinsessan (Disney/Egmont Kårnan AB) for young girls.](image)

Furthermore, referring to Mathes [43], impressions of other people’s personalities based on physical appearance not only persist but deepen over time. As a result, research lends credit to the folk-psychological notion that ‘first impressions are lasting’. Given that people seem to fall back upon social strategies when interacting with virtual agents in ways similar to how they interact with human beings [55], it is likely that the principle ‘first impressions are lasting’ has some bearing for virtual agents as well.

An implication is that if visual aspects are not carefully considered and articulated in research and development on virtual 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. [26]. A virtual 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 the system was going to be male lawyers. However, after realizing that the lawyers’ (female) secretaries were the ones who most frequently used the system, they became aware that the appearance and behaviour of the character disturbed these users. Thus, they designed a new character, with 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 appearance – the physical personality – also had to be redesigned.

Explanations for the neglect of visual design issues

When discussing the ‘neglect’ of visual design issues, it is important to emphasize that it is restricted to the research areas of virtual agents in general and virtual pedagogical agents in particular.

Outside these research areas, articles and discussions can be found, for example, of the representation and expression of gender in games and Internet based web presenters (e.g. webbots or chatbots). However, there is little or no dialogue between these areas and agent research – and furthermore, the very special context of human-agent interaction in pedagogical settings is an unexplored territory.

Computational constraints

One possible explanation to this neglected in virtual agent research is that visual design cannot be readily approached with deductive methodologies. Much of the research on virtual agents is computational oriented. The work on gestures, facial expressions and synthetic speech largely involves the development and refinement of computational models and algorithms. Once developed, speech and gestures may be easily produced and reproduced by such algorithms.

With respect to this, the visual design is all another matter. The catch is that even if gestures, facial expressions and, in particular, speech have to be individualized for different virtual agents (which can be accomplished
by adjusting parameters in computational algorithms) – the basic model for each of these aspects is more or less stable over time. The visual design, on the other hand, has to be related to rapid and unpredictable changing trends in graphical styles and fashion (affecting dressing, haircut, glasses, etc.). The inevitable demand for a certain degree of novelty in a trend sensitive business is a particularly important, but totally non-computational, aspect in most of today’s visually related design areas.

**Cultural constraints**

Another possible explanation for the neglect of visual design aspects in virtual agents may be that the influence of appearance 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 [21] [30] [36]. Unattractive people are considered as less socially competent, less willing to co-operate, more dishonest, unintelligent, psychologically unstable and antisocial [30] [32] [36] [49]. Negative reactions to ‘unattractive’ people are also more severe. For instance, Ahola [73], 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 those who were. Furthermore the crimes were perceived as less forbidding or repugnant if the perpetrator was good-looking.

Many people may instinctively disregard such findings – especially when they imply that we draw conclusions and pass judgement on appearance even in such a ‘rational’ domain as the legal one, as they challenge fundamental ethical and humanistic principles of a democratic society. In folklore, there are also proverbs that warn us against doing so: ‘Don’t judge a book by its cover.’

**Conclusion and discussion: Research question # 1**

To conclude and answer research question number 1: The visual appearance or representation of virtual pedagogical agents does matter – not least from a pedagogical point of view. In this sense, all traditional visual aspects of character design (body, face, hair, clothing and attributes) must be regarded as important.
There is also a definitive need for more systematic research concerning visual issues – not in order to establish general guidelines (which must be regarded as impossible), but to shed light on the possibilities and consequences of different visual design issues and to highlight important relations and phenomena. This could, for example, be a basis for design support in the form of checklists that design teams could relate to during the development of a virtual pedagogical agent.

It should also be emphasized that this thesis does not advocate any superior importance for visual issues over other agent related issues. Without the hitherto impressive work invested in the development of models and algorithms for artificial intelligence, knowledge systems, user models, personality, emotions, speech technology, gestures, and so on – there would be no visual design issues to bother about. But given these impressive achievements, it is about time to address a more holistic approach which also addresses visual design. For example, a clever design as to visual aspects may clarify which computational problems are most important from a pragmatic perspective in order to have a functional pedagogical agent system that can be used in a class room.
CHAPTER 6: VISUAL ASPECTS AND PEDAGOGICAL ISSUES

Paper II ‘Social and Visual Style in Virtual Pedagogical Agents’ addresses research question number 2: ‘In order to give an example of a visual aspect in relation to an educational context: Is there some kind of relation between visual characteristics of virtual pedagogical agents and characteristics of the learner?’

This is actually a story that goes several years back in time. In 2001, a set of new test materials, to be used by master students in cognitive science, was prepared. The test was targeted to investigate user attitudes towards the use of virtual pedagogical agents. Initially, the focus was on the agent’s social/communicative style, but in the process of designing a general and neutral agent, the suggestion of variation in the visual (graphical) representation came up. This line of thoughts was then related to reasoning presented in comic artist Scott McCloud’s seminal book ‘Understanding Comics: The Invisible Art’ [46], resulting in the idea to combine a social/communicative style test with a visual/graphical representation test. Three years later, a second, more extensive study was run in collaboration with a cognitive science master student.

These two studies constitute the basis for Paper II and are summarized in this chapter.

Iconization and subjective involvement

Comic artist Scott McCloud has proposed a theory regarding the iconic style used in many comics and the possibility for the reader to involve themselves into the characters [46]. The underlying mechanism consists of the image and concept of oneself being highly iconic (Figure 5).

When people interact they usually see the features of the other in vivid detail, but they also sustain a constant awareness of their own faces, and this mental image is highly iconic. Therefore, subjective identification and social affinity with an iconic character requires less effort compared to a realistic character, the latter taking the role of an object (another person). (NB: In the thesis, the expression ‘realistic vs. iconized’ is interchangeable with ‘naturalistic vs. stylized’.)
The two studies

Paper II presents both the original *social/communicative vs. visual style* study and the later and more extensive study mentioned.

**Study 1**

The first study used three different paper sheet demonstrators, picturing four sequential steps of an imaginary pedagogical multimedia program (Figure 6).
Each sheet presented one of three pedagogical agents differing in degrees of realism versus iconicity.

The agents had the role of assisting in the solving of a brain puzzle quiz. The demonstrators were used in the first part of the study as the point of departure for choice tests and interviews regarding preferences, responses and attitudes towards variations of the agent with respect to: (i) social/communicative style (relation and task oriented vs. strictly task oriented); and (ii) visual/graphical style (detailed/realistic – semi detailed/cartoon – simple/iconized). In the second part of the study, the participants completed a learning style inventory.

**Study 2**

The second test presented 90 elementary school students, age 12-15, with an introduction to a prototype of an educational game. In the educational game, the student was to take the role of a journalist going to Istanbul (Turkey) to cover a story.

![Example screen dump of the agents used in study 2.](image)

In the experiment, the student was randomly presented with one out of two possible settings:

(1) An *Instructor Version*, where the student was told that there was a chief
editor in London who would be his or her instructor. The chief editor would formulate the missions, orient the journalist (student) and provide necessary information at critical stages. The journalist (student) was to report back to the chief editor who would evaluate the reports and tell what was well done and what needed more work.

(2) A Companion Version, where the student was told that there would be a companion journalist with whom s/he would conduct the missions. The student was also told that it was important to co-operate with the companion who although not completely reliable when it came to knowledge, had some of the keys necessary to complete the missions.

The student was then presented with a set of eight pre-scripted, randomly positioned agents and asked to pick one of them as instructor or companion, depending on the scenario presented during the introduction. The eight agents consisted of two males and two females. Each was rendered in two versions: 3D and a more cartoon-like 2D (Figure 7).

Next, an interview took place where the student first was asked to motivate his/her choice of instructor/companion. Thereafter, the student was asked to choose between a social and task-oriented versus strictly task-oriented instructor/companion, and the student was also asked to motivate his/her choice. The study was then finished with the student completing two learning style inventories.

Results and conclusion: Research question # 2

The second research question: ‘Is there some kind of relation between visual characteristics of virtual pedagogical agents and characteristics of the learner?’ was affirmative. In particular, the two studies showed a correlation between a preference for an iconized visualization and for a relation oriented, more ‘subjective’, agent.

The results of the two studies also provide some support for the other aspects of research question number 2, but further investigations are required in order to better validate as well as to understand the results.

A tentative conclusion with respect to design is that if the goal is to design a pedagogical agent, rich in subjectivity and more relationally oriented, icon (stylized) visualization may be the better choice. Likewise, if the goal is to design an objective, task oriented pedagogical agent, a realistic representation may fit better.
Another outcome of the studies regards user variability. The studies showed considerable variability in user preferences, where different users strongly advocated their preferences as to both visual and social style. As regards social style (relation oriented vs. task oriented) this can be related to Bickmore’s [16] observations that the agent REA in the social condition evoked strong and diverging reactions. Likewise, Bickmore’s studies of the health advisor agent Laura [16] indicate that user appreciation of her relational styles ranged widely.
CHAPTER 7: VISUAL STEREOTYPES AND VIRTUAL PEDAGOGICAL AGENTS

Paper III ‘Aesthetic Stereotypes and Virtual Pedagogical Agents’ addresses research question number 3: ‘If the answers to questions (1) and (2) are affirmative, what are the consequences of, given the strong impact of visual stereotypes in the society?’

The paper is written in an argumentative tradition – disclosing and reflecting on the concept of visual stereotypes, the relation to virtual pedagogical agents, the new context brought on by the agents, and possible future effects of stereotyping in relation to the new and hitherto unexplored interactive human-agent qualities.

Visual aspects

So far, it has been shown that: (1) it is highly plausible that visual (graphical) aspects in virtual pedagogical agents have an impact; and (2) visual and social characteristics of virtual pedagogical agents may be significantly related. In this chapter, the consequence of these findings will be discussed in the light of visual stereotypes. The aim is to clarify aspects and connections that may be important in the generation of visual design criteria for virtual pedagogical agents in relation to ethical, social and political standpoints.

As already discussed, it is likely that visual design issues will be of importance for future agent-based educational software – not least when it comes to ‘first glance’ effects and promotion. A well-designed agent-based educational program as to pedagogical strategies, knowledge competence, gestures, animation and speech-technology may fall short if the interface is poorly designed in a graphical sense. To cite a manager from a Swedish package company, ‘When the content is more important then the package – the package is important.’ Even so – a careful and competent design of the visual instances will probably also be essential for smooth and effective communication (interaction) as well as for the ambition to engage and motivate the learner. These line of reasoning have additional support in experiments, were aesthetic factors have been shown to have an impact on cognitive tasks [52] [66]. There is also a growing awareness in the field
of HCI/Interaction Design that aesthetic factors may add an important contribution to the experience [16].

Visual stereotypes

In order to address some of these issues, the concept of visual stereotypes (in relation to virtual pedagogical agents) may be fruitful. This approach also addresses the questions of to what degree and in what circumstances visual issues have an impact regarding virtual pedagogical agents.

Visual properties

In human-human social interaction the visual perception of other people is known to play a central role, with profound effects on our attitudes as well as behaviour. In order to discuss visual stereotypes in relation to visual design issues in virtual agents, the basic properties of visual stereotypes may be differentiated according to: (i) dynamic visual properties such as gestures, facial expressions and gaze (which, as already shown, are extensively researched within the agent community); and (ii) static visual properties such as body and face properties, skin, hair, clothes and attributes. In the following discussion, these static properties are considered to be the basic source for the construction of visual stereotypes.

Cognitive implications of visual stereotypes

![Figure 8: Examples of visual stereotypes](image)

1a 1b 1c 1d
Figure 8 presents four examples of visual stereotypes. Many observers will see in these pictures: (1a) a teenager, (1b) a housewife, (1c) a craftsman, and (1d) an air hostess. A visual stereotype, in this sense of the term, consists of a number of static visual attributes in a person that will make a majority of observers perceive the person as an illustration, or a typical instance, of a human group, a professional group, or a social group. The visual input, thus, activates expectations on other aspects of the person: how s/he is likely to behave and to talk; what s/he can be expected to say or not say; what attitudes and opinions s/he will be likely to have, and so on. In this way visual cues carry social baggage.

As described above, stereotypes appear as a negative ingredient of human behaviour, but it is not that simple. From a cognitive perspective, the use of stereotypes is absolutely necessary in order to interact with other people. They function as cognitive short cuts for making action and life tractable for human beings. Instead of becoming overloaded by analyzes, thoughts and questions when encountering unknown people, their visual appearance helps to quickly situate them in order to focus on interaction as such [21]. In this way, visual stereotypes frame peoples expectations and are used for building common references in conversations about other people. In brief, (visual) stereotypes are a navigation tool in a social environment that would otherwise be overwhelmingly complex and demand a practically insurmountable burden of cognitive processing [61].

Visual stereotypes and virtual pedagogical agents

As described by Reeves and Nass [54], social behaviour and strategies from the real world are often reproduced in interaction with computers. An extension of these findings is that visual stereotypes in the real world may be reproduced in the interaction with virtual characters.

The conclusion is that positive as well as negative aspects of real-life visual stereotypes reappear in interactive media with their invaluable function as cognitive tools for handling a complex social environment, as well as their problematic normative function.

Novel possibilities …

There are novel possibilities introduced by visual stereotypes in virtual pedagogical agents. Elaborating on the possibility to break with, or exploit, (visual) stereotypes for pedagogical purposes, an interesting oppor-
tunity is to enable the exploration of a broad range of identities and to extend possibilities for social identification and role modelling.

... and risks

As to risks, the construction and promotion of idealized super people with ‘perfect’ bodies and looks (and even lives) has abounded in non-interactive media such as TV, video and magazines for a long time. (Figure 9).

![Figure 9: Stereotypes from ‘traditional’ media: Tarzan (© Burne Hogarth); Sean Connery (Film icon); Paladin avatar (Game character, World of Warcraft); P’Gell (© Will Eisner); Jane Russel (Hollywood star); Lara Croft (Game heroine, © Core Design/Eidos)](image)

This portrayal of the ideal can be taken *one step further* with interactive computer media. A key difference lies in what is otherwise seen as a central potential of virtual characters – namely their *interactivity*, relying on ‘autonomous’ algorithms and models for behaviour, personality, emotions, social strategies and memory. There is a plausible risk is that the distance between users or learners and those ‘ideal super people’ is diminished. And in an era already desperately pursuing perfection in appearance, this might have detrimental effects on people’s self-image and self-esteem.

**Summary: Research question # 3**

Positive and negative sides in the use of visual stereotypes known from real life interaction reappear in interactive media. This has been exemplified by: (i) their positive and invaluable function as cognitive tools for
handling a complex social environment, and (ii) their negative and problematic normative function that can make what diverges from a visual stereotype be perceived as odd or abnormal.

There are novel possibilities and risks introduced by visual stereotypes in virtual agents. Among the possibilities is the opportunity to break with stereotypes for pedagogical purposes to enable an exploration of a broad range of identities and extend the possibilities for social identification and role modelling. As to novel risks, the introduction of virtual agents may have detrimental outcomes on peoples self identification and self image. This could happen due to the interaction with visually idealized stereotypes with communicative, social and emotional qualities – an extended interactivity not found in media characters this far.
CHAPTER 8: SUMMARY OF THE THREE INCLUDED PAPERS

The extracts of Paper I, II & III presented in the three preceding chapters represents a selected summarization. The key results of the three included papers (Paper I, II & III) are summarized below.

(NB: The papers themselves provide much more elaborated discussions on these topics as well as additional themes.)

Summary

(1) It is highly probable that visual (graphical) aspects in virtual pedagogical agents have an impact on learners in terms of pedagogical effects and outcomes.

(2) There are significant relations between visual and social characteristics of virtual pedagogical agents and social/communicative characteristics of learners. Specifically, a relationship has been proved between learners’ preferences as to: (i) the visual representation of the agent in terms of naturalism vs. stylization; (ii) the social/communicative style of the agent in terms of strictly task orientation vs. task- and relation orientation; and (iii) the social/communicative style of the learner. There is also evidence that user groups differ in characteristic ways. This indicates a possibility to define pragmatic and relevant categories of user groups when including virtual pedagogical agents with different alternative visual styles.

(3) With respect to visual stereotypes, virtual pedagogical agents: (i) may reproduce stereotypes from everyday real life human-human interaction, as well as from traditional visual media; and (ii) the introduction of virtual pedagogical agents with their extended interactivity may even strengthen certain detrimental effects that idealized stereotypical instances have on society and people’s self-esteem.
DISCUSSION

&

FUTURE RESEARCH AGENDA
CHAPTER 9: DISCUSSION & FUTURE POSSIBILITIES

As argued in this thesis, there is strong evidence that visual aspects have an impact on the interactive experience of virtual (pedagogical) agents. Furthermore, it is possible to identify user groups sufficiently large and well-defined to benefit from different visual renderings of virtual pedagogical agents. Acknowledging the importance of visual aspects inevitably puts the light on ethical questions as the concept of stereotyping.

The thesis has already included extensive parts of discussion. Notwithstanding, there are additional aspects of virtual pedagogical agents touched upon in the included papers, as well as in other work, that point towards interesting areas of future research.

The visual design space of virtual agents

Paper I took a closer look at different aspects in the design of virtual pedagogical agents, such as body, face, hair, clothing, gestures and facial gestures. The paper also discussed the dimension of naturalism – stylization (also expressed as ‘realism – iconization’ in this thesis). In a recently presented paper [30] these lines of reasoning have resulted in a more elaborated conceptualization of static and dynamic visual properties.

As mentioned in this thesis, dynamic properties, being accessible for a computational approach, have been extensively researched as to virtual agents. They are also explicitly discussed in the domain of animation, as in the extensive canonical work of Frank Thomas and Ollie Johnston, The Illusion of Life: Disney Animation [65].

When it comes to static properties, on the other hand, empirical research is rare. Only recently a few papers in the area of agent-research have touched upon different dimensions of the static visual properties of virtual (pedagogical) agents:

– Churchill et al. [25] [26] claim that the design of character appearance is central to charting out which design issues ought to be explored for embodied conversational characters. The authors explicitly mention the Look and propose five design dimensions of look to be explored: the degree of humanoidness; the degree or stableness versus changeability
in appearance (e.g. morphing); the degree of animation versus static-
ity (i.e. the extent to which the character can move); 2D or 3D visual
rendering; and degree of realism (from photo realism to line drawing).
However, in the actual development of their own character Will, no
exploration of these design dimensions is reported.

- Ruttkay et al. [57] discuss how to evaluate embodied conversational
agents (ECAs). Among much others, they present a short paragraph
with a few visual parameters under the label ‘Look’: personification
(person, creature/animal, or non-living object); realism (realistic, art-
istic, or cartoon-like), dimensions (2D or 3D), physical details (head,
head + neck, torso, or full body).

- Gratch et al. [29] have emphasized the importance of visually stylized
presentation styles in a seminar report. Alas the report this far only has
resulted in an extended abstract.

Character by Design [32], is the first to take a firmer grip on design is-
ues regarding the conceptualization of virtual characters – however,
a transformation from game design to a pedagogical context is not
straight forward.

**Visual static dimensions**

Among the possible visual static dimensions, ‘realism vs. iconization’ is
by far the most discussed, and it is also the dimension that has been most
studied, e.g. [11] [29] [68] [69]. It is obvious, though, that the (static)
visual design space is far more complex than just the simple realism –
iconization dimension, and a more elaborated proposal for a virtual agent
related design space is presented below and in Figure 10 [30]:

- **Humanness**: An agent can basically be modelled upon humans, animals
(or other creatures), and non-living objects – or some combination
of these three entities, exemplified in Figure 10 by three well-known
agents: Steve (© CARTE, University of Southern California), Herman
the Bug (© IntelliMedia, North Carolina State University), and the Mi-
crosoft Office Assistant ‘Clippit’ (© Microsoft Corporation).

- **Basic ‘physical’ properties (shape and colour)**: For a humanoid agent, for
example, there are basic ‘physical’ properties, such as body-type, face,
colour and type of skin, hair, clothes and various attributes.
– **Graphical style**: The graphical style of an agent’s appearance relates extensively to artistic qualities, and there are several ways to describe these properties. For the purpose of virtual agents, two dimensions are of particular interest:

**Naturalism vs. stylization**: This dimension is complex, with the stylized alternatives, in particular, spanning over a wide range of different styles or expressions. The diversity can be exemplified by considering a character from *The Sims* representing near naturalism, a Picasso-styled face, and a *Peanuts* inspired face, representing three different stylized expressions. Note that there is no simple, linear relations in the design space of naturalism – stylization.

**Detail vs. simplification**: Referring to Figure 10, lower left, the leftmost character can be referred to as detailed. By means of reduction it may be transformed into a Disney-styled figure. Further reduction turns it into a very simplified (and stylized) cartoon. It is also possible to start with a photo, as being detailed, and then adjust it into an image with only three colour levels: white, grey and black. This new picture would still be ‘realistic’, but less detailed.

In the above presentation the 2D vs. 3D dimension is deliberately left out as a dimension in its own right. The argument is that: (i) 2D and 3D can be handled as different styles or expressions within the ‘Graphical style’
dimension; and (ii) 3D is basically to interact in a 2D-simulation of 3 dimensions (unless perhaps when using virtual reality equipment like helmets and gloves). It is also possible to inhabit a 3D-environment with flat 2-dimensional figures and scenes – looking like thin lines if viewed from the side. This is actually a visual joke that comes to use in film comedies, e.g. *Who Framed Roger Rabbit* by Touchstone Pictures [71].

**A criticism of user evaluations in the domain of virtual (pedagogical) agents**

As has been argued throughout this thesis, visual aspects do matter. Likewise it has been showed that different identifiable user groups may vary significantly in their preferences as to the visual style of virtual (pedagogical) characters. An implication of this is that the results of many user evaluations carried out within the virtual agent research domain are not possible to generalize (to the extent they have been generalized). Typically, one can basically only state that the outcome of a user evaluation tells that *one specific instance* of a virtual agent was more appreciated than *another specific instance*. Consequently, it is hardly possible to state that, for example, realistic characters are ‘better’ than cartoon-like characters after a comparison of two or three specific representations, singled out from the huge and complex visual design space of virtual characters.

Furthermore, the ignorance of possible significant variations between groups of users, leads to the misleading concept of ‘the User’. This approach – treating users as a homogenous and standardized group – is not seldom applied to evaluations of virtual agents. Accordingly, the results of several evaluations fail to reveal possible relevant and significant correlations with regard to specific groups of users.

On the other hand, by being aware and precise about visual parameters, and by identifying significant user groups – user evaluations may be more generalizable and the results more interesting as to relevant correlations between different agent parameters and user characteristics.

**Virtual ‘assistants’ in the health domain**

In the near future, the most interesting and fruitful area for virtual pedagogical agents is perhaps the health care sector. There are several reasons for this, two of them being: (i) the long existing problem with motivat-
ing people to follow health or rehabilitation recommendations once back home, when they only meet their therapist or physician once a month and in the meantime are supposed to rely on written instructions; and (ii) the forthcoming demographic shift, leaving an increasing number of older people with a relatively decreasing number of therapist and physicians together with an uncertainty as to financial resources.

In the light of this, there is an interesting possibility to improve the quality of self care and rehabilitation in the home through the support of agent-based systems. In this context, the social and emotional qualities of virtual pedagogical agents can motivate and engage substantially more users or patients to keep up with recommendations and rehabilitation programs. The most interesting work in this area is being done by Timothy Bickmore at *Northeastern University*, e.g. [17] [18] [19].

There are also some recent research programs studying the possibilities of virtual therapists as the *Centre for Spoken Language Research* (CSLR), University of Colorado.

It should be emphasized that ‘virtual therapists’ are not to replace real human therapists: (i) virtual therapists are not intelligent or sensitive enough; (ii) there are ethical and legal complications hardly even thought of; (iii) a virtual therapist is actually only a tool for the therapist to improve the outcomes of home-based rehabilitation programs; and (iv) even with the addition of virtual therapist, there will most probably be a strong connection between the engagement and amount of invested time on behalf of the real human therapists and the result of a rehabilitation program.

In other words, a virtual therapist is not a ‘solution’ that will replace human therapists – but they may be a useful tool for therapists to improve the outcome of rehabilitation programs.

**Virtual agents as a research tool**

A more recent track of virtual characters is their potential as a research tool in other research domains as psychology, cognitive sciences, neurobiology, anthropology, cultural studies, etc.

Given the evidence that responses and behaviour in human-human interaction are transferred into human-agent interaction, it may be possible to repeat many studies involving humans. As a virtual character basically is a well-defined set of data, it is possible to systematically manipulate separate variables (or sets of variables) in the virtual character. For example,
it is possibly to control variables such as: the colour of the hair, the shape of the cheek, the gender, the age, the dressing, the pitch and/or tempo in the speech, the way of phrasing, the sweeping in a gesture, the intensity in the gaze and the way to instruct (just to mention a few).

In other words, there is a potential to study different socio-cognitive and socio-psychological phenomenon with a calibrated and well defined method – something that hitherto not has been possible with real humans. The possibilities as well as the consequences of such studies are, to say the least, speculative.

Future research agenda

As this thesis has been more or less continuously discussing different themes around virtual (pedagogical) agents, there are also several research agendas implicitly presented throughout this thesis. Some of these suggestions are recapitulated in short below:

(1) There is an urgent need to further explore visually related aspects of virtual (pedagogical) agents in order to handle the outcomes of user evaluations – or else results will continue to be blurred by the uncontrolled impact of different visually related variables.

(2) It will be fruitful to attempt to identify relevant and identifiable user groups, in order to discover both trends and significant correlations between agent characteristics and user (group) characteristics.

(3) The rise of virtual (pedagogical) agents that lurks around the corner will provoke complex and not even thought of ethical and legal complications.

(4) The potential to use virtual agents as research tools is extremely interesting. Especially within domains such as psychology, cognitive sciences, neurobiology and social sciences, the use of virtual agents to control experimental variables may have a considerable impact.

To summarize: the research field of virtual (pedagogical) agents is immense.
REFERENCES

VPAs, CALs & Articles
REFERENCES

Virtual pedagogical agents, VPAs (Page 3)

*AutoTutor:* Cognitive Science Lab (CSL), Institute for Intelligent Systems, University of Memphis, Memphis, TN. http://www.autotutor.org/

*Herman the Bug:* IntelliMedia, North Carolina State University, Raleigh, NC. http://research.csc.ncsu.edu/intellimedia/


*Steve:* CARTE, University of Southern California, Marina del Rey, CA. http://www.isi.edu/isd/carte/


*FearNot:* VICTEC / eCIRCUS http://www.macs.hw.ac.uk/EcircusWeb/

*Olga:* CID NADA & TMH (KTH), Linguistics (SU), SICS & Nordvis http://www.nada.kth.se/~osu/olga/


Computer assisted learning systems, CALs

1. **PLATO** (Named by the Greek philosopher Plato). One of the first generalized computer assisted instruction systems and originally developed by University of Illinois in the 1960s.
   PLATO IV from 1975 displayed a new interface with monitor and keyboard. Economically a failure – PLATO nevertheless pioneered key concepts such as on-line forums and message boards, on-line testing, e-mail, chat rooms, picture languages, instant messaging, remote screen sharing, and multiplayer on-line games.

2. **TICCIT** (Time-shared, Interactive, Computer-Controlled Information Television). First developed by the MITRE Corporation in 1968 as an interactive cable television system.
   Trial implementation as a computer-assisted instruction system for English and algebra courseware 1975-77. Commercialized by the Hazeltine Corporation.

3. **Virtual Speech Therapy Systems** [LSVT™-VT, Sentactics™-VT, ORLA™-VT, C-Costa™-VT]: Center for Spoken Language Research (CSLR), University of Colorado, Boulder, CO.

4. **Laura**: Timothy Bickmore, Northeastern University, Boston, MA.
   http://www.ccs.neu.edu/home/bickmore/ (2006-10-20)

5. **FearNot**: VICTEC / eCIRCUS
   http://www.macs.hw.ac.uk/EcircusWeb/ (2006-10-20)

Articles


PAPER I:

DESIGN OF ANIMATED PEDAGOGICAL AGENTS – A LOOK AT THEIR LOOK

Design of animated pedagogical agents—A look at their look

Agneta Gulz\textsuperscript{a}, Magnus Haake\textsuperscript{b,*}

\textsuperscript{a}Division of Cognitive Science, Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden

\textsuperscript{b}Department 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.

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 choose 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 Thörisson, 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). An addition of animated agents to intelligent tutoring systems can, in other words, be seen as an attempt to fulfill 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 (McCauley 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 (Ovatti and Adams, 2000); Herman the Bug (Lester et al., 1997); Talking Head (Moundridou and Virvou, 2002); Laura (Bickmore, 2003). In other systems the visualized agent primarily acts as a learning companion—e.g. Trouble Maker (Almeur et al., 1997) and Steve3 (Rickel 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; Koda 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., 1999).

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.

The intelligent tutoring system that the animated pedagogical agent is based upon is assumed; we are thus not discussing design features such as level of competence of agents.
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 (Bécheiraz 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 (Isbister 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 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 represented 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. 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, unproblematised 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 environments.

---

The visual characters used in the study were four predefined characters from the Microsoft Agent package (Microsoft 1999).

The process may not be verbally articulated but well documented in the form of character descriptions and ‘model sketches’ (Thomas and Johnston, 1984).
conversational characters. The authors explicitly mention the design of the Look 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.

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...
involves computation. For instance, some of the research on design of gestures for agents is based on ethnomorphic 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.13

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.14 Furthermore the crimes were perceived as less forbidding or repugnant if the perpetrator was good-looking. However, many people may find this result—that we draw conclusions and judge on appearance, even in such a rational domain as the legal one—both surprising and unwelcome. Such behaviour is not sanctioned by ethical and humanistic principles. In folklore there are proverbs that warn us against doing so: “Don’t judge a book by its cover”; “Don’t judge a dog by its hair.”15

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 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-

tions of appearance and observable physical cues profoundly affect both beliefs and behaviour (Milord, 1978; Schneider et al., 1979; McArthur, 1992; 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 small16 (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 [...]

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 hundet efter hären!)

16Even though, of course, some voice characteristics are also quickly picked up.
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—on a 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

---

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—such 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
declare that ‘‘[t]he value of the costume in creating a
personality cannot be overestimated’’ (Thomas and John-

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 stereo-
types 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. Further-
more, it has repeatedly been shown that we draw inferences
about people based on the clothes they wear (e.g. Bardlak
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

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

Fig. 6. The morphological characteristics that mark a baby’s face are
large eyes relative to the rest of the face, fine, high eyebrows, light skin and
hair colour, red lips that are proportionally larger, a small, wide nose with
a concave bridge, and a small chin. The facial features are also placed
lower on the face (Branham, 2001). All of these characteristics can
regularly be found in the stereotypical portraying of females in comics and
animated film. Also note the stereotypical forward bent, with eyes looking
up (Thomas and Johnston, 1984)—a pose frequently used when females
intersect with their male counterparts. Prototypical drawing by Magnus
Haake based on the different characters in a Swedish Disney Princess
magazine for young girls (Prinsessan, Nr 1, 2005). © Magnus Haake.

...
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, 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.
But even though explicit discussions on drawing and visual style are rare, much is implicitly expressed on the topics. Both in Thomas and Johnston (1984) and in the computer game industry (Gamasutra, 2004), one can notice a continuous experimentation with visual aspects and corresponding responses from the audience.

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 engagement (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. 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, the area of computer games is one that is repeatedly advised not to start out with any programming before the visual aspects and the story are thoroughly worked out. 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 academies 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—and without

---

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

24In 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 stereotypical 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.25 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


25Within the area of animated pedagogical agents, it seems that there are design teams with scientists/technical expertise and/ or graphical designers/animators—but there are no reports of the design space explored in this part of the design process.


PAPER II:

SOCIAL AND VISUAL STYLE IN VIRTUAL PEDAGOGICAL AGENTS

ABSTRACT
The paper addresses aspects of virtual pedagogical agents’ visual style (realism – iconization) in relation to their social style (task oriented – relation oriented). Two studies are presented that investigate which visual and social styles users prefer and how they articulate their preferences. The first study involved 42 university students; the second study involved 90 elementary school children. Special emphasis was put upon two hypotheses, grounded in cognitive theory: (i) iconicized visualization may be better suited for representing a relation oriented, subjective agent – and therefore preferred by users who prefer a relation oriented agent; (ii) realistic visualization may be better suited for representing a task oriented, objective agent – and therefore preferred by users who prefer a task oriented agent.

The results of the two studies provide some support to these hypotheses. Cognitive theories are exploited to interpret the results, and possible design considerations are discussed.

Keywords
Agent, pedagogical agent, visual style interface, social style interface, pedagogical role, individual cognitive differences, motivation, human-computer interaction, user studies.

1. INTRODUCTION
An aspect that has been surprisingly little researched regarding virtual pedagogical agents is their visual appearance in terms of facial shape, body and costume, visual style, etc. Yet, the evidence that many basic patterns from human-human interaction recur in human-computer interaction [12] gives us reason to address visual rendering issues in research on animated agents. It is well established within social psychology that aspects of visual appearance have considerable impact on how people access other people. We are profoundly affected — in terms of behavioral responses as well as beliefs and attitudes — by others’ body shape, facial looks and clothing. Furthermore, research supports the notion that “first impressions are lasting”, i.e. impressions of someone’s personality based on physical appearance, not only persist but also deepen over time [8]. This principle is likely to hold for animated agents as well (cf. [6]).

One implication of this line of reasoning is that if visual appearance of animated agents is ignored, rather than being carefully considered and articulated in research and development, agents may fall short of motivating and engaging users. The survey presented in the article Design of Animated Pedagogical Agents – a Look at their Look [5] shows how visual rendering issues have been neglected in research on animated pedagogical agents. The main thesis of the paper is that users’ visuo-aesthetic experience of animated pedagogical agents is too important to be disregarded, with respect to the goals to motivate and engage.

The present paper addresses the issue of visual style (realism – iconization) in relation to the notion of an agent’s social style (task oriented – relation oriented). In what follows, we introduce the topic and present two empirical studies.

2. VISUAL AND SOCIAL STYLE
2.1 Visual Style: Iconicity vs. Realism
Virtual pedagogical agents can be visualized in a number of ways. In the present paper, we investigate two different visual styles: realistic and iconic. Here realistic signifies a rather realistic 3D-rendering compared to iconic, which signifies a somewhat simplified character of the type to be found in Marvel Comics, for example. Various theoreticians have presented ideas on realistic versus iconic agents with respect to user involvement. Welch et al. [13] argue that pictorial realism increases involvement and the sense of presence in a digital environment and even is a condition for human cooperation with an animated agent.

Moving to another domain, McCloud in his seminal book Understanding Comics [10] argues to the contrary that audience involvement is increased by iconization. The underlying mechanism consists of the image and concept of oneself being highly iconic. When people interact they usually see the features of the other in vivid detail, but we also sustain a constant awareness of our own faces, and this mental image is highly iconic. Therefore, subjective identification and social affinity with an iconic character requires less effort compared to a realistic character, the latter taking the role of an object (another person).

All in all, we find diverging claims regarding realism versus iconicity with respect to the ability to involve and engage, even if there is not much of an explicit discussion. At the same time, we have the impression that there is a non-reflected assumption within the domain of virtual pedagogical agents that a visually realistic agent is an obvious goal to strive for. In any case, there is a need for systematic studies.

2.2 Social Style: Task Oriented vs. Relation Oriented
2.2.1 The Social Computer
As Reeves and Nass have demonstrated [12], people respond to computers in social ways and apply behavioral patterns of politeness, flattering, etc. These effects become even stronger when the
computer artifact is an anthropomorphic character, i.e. a character with humanlike shape and other humanlike properties. Several researchers emphasize that a close emulation of the features that are present in human-human face-to-face communication is crucial in order to obtain positive results with pedagogical agents, as such emulation contributes to smoother communication and makes the interaction more stimulating, motivating or engaging (e.g. [4]; [9]).

2.2.2 Different communication styles
Yet human beings do not all interact and communicate in the same manner, and we can speak of different communicative styles (e.g. [11]). The communicative style dimension addressed in this work is that of (1) a strictly task oriented communicative style versus (2) a combined task and relation oriented communicative style. Given a pedagogical context, the two styles can briefly be exemplified as follows: (1) a coach (instructor, teacher, learning companion, etc.) who is strictly task oriented, sticks closely to the task, provides information in a succinct and objective way and focuses on factuality, and (2) a coach (instructor, teacher, learning companion, etc.) who is relation oriented (as well as task oriented) will, apart from contributing to the solving of the learning task, also work on developing of a social relationship with the learner; personalizing the task, being more subjective, and focusing less strictly on the task in the dialogue. The next section describes features of relation oriented communication in more detail.

2.2.3 Relation Oriented Communication
In human beings there are many features that contribute to relation oriented communication. A variety of non-verbal behavior is involved, such as forward lean, body and facial orientation, smiling, nodding, gaze and gesturing (cf. [2]). On the verbal side, the dialogue does not only regard the task at hand; it may contain small-talk, conversational storytelling, getting-acquainted-talk, joke-telling, sharing of personal experiences, preferences and opinions. Certain voice features are also often present such as greater warmth and expressiveness, reinforcing interjections such as “mm-hmmm” and more variation in pitch, amplitude, duration and tempo [2].

2.2.4 Implementations
Two of the most interesting implementations involving the features mentioned are the agents REA [2] and Laura [2]. Whereas REA has the role of a real estate agent who interviews potential home buyers and shows them around houses, Laura has a more professional manner with authoritative speech and provides accurate information in a succinct way. The animation is limited to deictic gestures. There is little expressivity and the agent shows no affect. The task and relation oriented Mentor agent works collaboratively with the learner. The goal is to demonstrate competence to the learner while simultaneously developing a social relationship to motivate the learner. Gestures incorporate both deictic and emotional expressions, and the agent shows various affects such as confusion, approval, excitement and pleasure. Sometimes the agent uses colloquial expressions, e.g. “What’s your gut feeling?” [1].

For a forthcoming study that we are currently planning, our goal is to implement as many as possible of the features described above in the design of a task oriented versus a task and relation oriented pedagogical agent. Dialogue aspects such as small-talk versus no small-talk, conversational storytelling and getting-acquainted-talk versus no such features, sharing of personal experiences and opinions versus no such features, and some of the vocalic behavior described above, will have first priority.

In our two studies reported on in this paper, the users, after having chosen their instructor/learning companion with respect to visual representation, encounter only descriptions of the two kinds of communicative styles, and are asked which one they would prefer: an instructor/learning companion that is strictly task oriented and whose talk only relates to the task, or an instructor/learning companion who apart from talking about the task, engages the user during pauses in small-talk, supplies personal information, and relates personal experiences and interests.

3. THE ISSUES
3.1 Study Motivations
Underlying the two studies presented in this paper as well as the forthcoming study is a desire to learn more about user/learner preferences regarding visual and social style and to learn more about possible reasons for the preferences. As Mc Cloud suggests (see subsection 2.1), identification, social affinity and the formation of relationships require less effort with an iconic agent than with a realistic agent. From this line of argument, we formed a hypothesis that a preference for a more subjective and relation oriented character will correlate with a preference for a more iconic character, and that a preference for a more objective and task oriented character will correlate with a preference for a more realistic character.

In the second study, we also included the variable of pedagogical role, hypothesizing that learner preferences regarding visual and social style would differ with the pedagogical roles of the agent (instructor vs. learning companion).

3.2 Study Questions
The following issues were addressed in the two studies:
(1) If a learning environment provides a set of animated pedagogical agents, with different degrees of realism – iconicity, to chose from: (1a) What do learners chose? (1b) How do learners articulate the motives for their choice?

1 Termed “task condition” versus “social condition” in REA, and “non-relational” versus “relational” condition in Laura.

2 There is also a third agent, the Motivator agent, who could be said to be only relation oriented.
(2) (2a) Do learners prefer (i) a task oriented pedagogical agent that “keeps to the learning task” in an objective manner or (ii) a task and relation oriented pedagogical agent that apart from being task oriented “socializes” with the learner in various ways? (2b) How do learners articulate their preference?

(3) Can any relationships be found between the results in (1) and (2) above? In particular, is there any support for the hypothesis that iconized visualization in agents is better suited for representing a relation oriented agent and that realistic visualization in agents is better suited for representing a task oriented agent?

In addition, in the second study:

(4) Does the role of the agent (instructor vs. learning companion) affect the learner’s choice with respect to visual style (realism – iconicity) and/or social style (task oriented – relation oriented)?

4. FIRST STUDY

The material consisted of three different Lo-Fi paper sheet prototypes, picturing four sequential steps of an imaginary pedagogical multimedia program. Each sheet presented one out of three pedagogical agents (differing in degrees of realism – iconicity) assisting in the solving of a brain puzzle quiz (Figure 1 below).

Participants were 42 university students, 23 women and 19 men, 19-25 years old. All had some familiarity with virtual pedagogical agents. After a brief introduction, the participant was asked to imagine an e-learning environment dealing with the basics of neuroscience, and was shown an example task. The participant was then told that the environment also supplied a virtual coach that gave feedback and problem solving advice, whereupon the three Lo-Fi sheets were placed in a circle (varying the positions between participants). The participant was now asked: “If you were going to work with this learning environment, which of the three characters would you prefer as your virtual coach?”. Next, the participant was asked: “Which agent would you prefer least?” After this, the participant was requested to motivate the choices.

Figure 1. To the left are the three different agents (differing in degree of realism – iconicity) used in study 1, and to the right is an example of one of the three Lo-Fi scenarios (with the middle form agent) used in the study.

---

3 Results relating to this issue will be reported elsewhere, as well as results regarding the relationship between learning styles and preferences as to social and visual style in agents.

4 The familiarity ranged from “being familiar with the MS Office paper clip” to “knowing a variety of play-and-learn-programs for children” to “having used several animated agent products, in database programs, simulation programs for education, etc.”.
After that, one of the experimental leaders verbally presented two scenarios: (i) one with a coach that focuses on the tasks and (ii) one with a coach with a richer and more complex social personality that, apart from advising and guiding, displays personal and social features. Finally, the participant completed a learning style inventory.

4.1 Results

4.1.1 Visual Style Preferences

Ten participants chose the most realistically drawn agent, 13 choose the middle form and 19 chose the most iconic. Female participants tended to a higher, but not significantly higher, degree than male participants to choose the most iconic agent.

Participants’ views on the advantages and disadvantages with realistic and iconic agents diverged. The most frequent theme was that of serious-childish, where the realistic agent was preferred as more serious and the iconized agent rejected as childish and not serious enough. Another frequent theme was that of distraction, used both in arguments for the iconized or middle form agent (seen as not so distracting), and in arguments against the realistic agent (seen as too distracting).

4.1.2 Social Style Preferences

The distribution of choices with respect to social style in agents was as follows: 12 (63%) of the men and 14 (61%) of the women was as follows: 12 (63%) of the men and 14 (61%) of the women.

Recurring arguments for choices are summarized in Table 1.

Table 1. Social style arguments (translated from Swedish).

<table>
<thead>
<tr>
<th>Social Style</th>
<th>Arguments (No. of arguments in parentheses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>For task oriented / (Against relation oriented)</td>
<td>“It may disturb the learning (11) “The agent should focus on my learning, everything else will disturb me.”; “It ought to stick to the subject matter and not talk of anything that does not have to do with the learning process.”</td>
</tr>
<tr>
<td></td>
<td>“It is not a human being (3) “I am not interested in getting to know an agent, it is no human being.”; “They are not living creatures; there should not be much about them.”</td>
</tr>
<tr>
<td></td>
<td>Not appreciating this kind of social relation (3) “I prefer my real friends.” “It would feel awkward. I don’t like the idea of a virtual human.”</td>
</tr>
</tbody>
</table>

| For relation oriented / (Against task oriented) | More confidence inspiring (6) “It would be more convincing.”; “You feel more confidence in the agent and learn better.” |
|                                                | More fun and pleasant (3) “It would have been great fun if it worked well.”; “It would be a stimulus, like having someone who cheers you up when it’s dull – the point in animation is to go one step further.” |
|                                                | More personal is better (2) “You learn better when it is more personal.” “It would be interesting; I’m always curious about humans’ inner lives.” |

In summary, participants’ views diverged on whether a relation oriented character is a good idea or not. The distraction theme that appeared regarding realistic visual style, also turned up with respect to relation oriented social style. The most frequent argument against a relation oriented agent was that it may disturb and distract from learning. Two other frequent and related themes regarded “agents being artificial agents and not humans” and “social relations with such an agent being weird or inferior”. The most frequently presented arguments in favor of a social agent were that it “increases confidence” and that it “makes it more fun”.

4.1.3 Visual and Social Style Preferences Together

Merging the results the following can be observed (see Figure 2): Among those who preferred a relation oriented agent, there was a significant tendency to choose an iconic agent ($\chi^2(2) = 6.50; p = 0.04$) whereas no significant tendency in agent choice could be seen among those who preferred a task oriented agent ($\chi^2(2) = 1.46; p = 0.48$).

5. SECOND STUDY

The second study pursued the issues of the first one. The two studies differed however in the following respects: (i) implementation: two versions of a computer-based pedagogical multimedia dummy in the second study vs. a Lo-Fi paper prototype in the first (Figure 3); (ii) age of participants: school children in the second study vs. university students in the first; (iii) learning subject: geography/social science in the second study vs. neuroscience in the first; (iv) number of agents: four different agents in two degrees of iconization in the second study vs. one agent in three degrees of iconization in the first; (v) addressing new issues on pedagogical role in the second study.

The difference (iv) was motivated by our desire to study agent choices where the iconicity-realism dimension was not too obvious. With the three agents in the first study, the iconicity-realism dimension was quite obvious to the participants. This probably affected the cognitive processes in choosing and analysing one’s choice. With the eight agents used in Study 2, the dimension was concealed to a larger extent.
5.1 Method

5.1.1 Participants

Ninety students, 48 girls and 42 boys (age 12-16), from 9 different school classes in Swedish elementary schools, participated in the study.

5.1.2 Materials

Two dummy versions of a scenario-based multimedia program for elementary school were developed for the study. In both versions, the student is to take the role of a journalist at a magazine, being sent to European countries to do article research. In the Instructor Version the student is guided by a virtual instructor and in the Companion Version accompanied by a virtual companion. Both dummies, created in Macromedia Director, include (i) an introduction where the program and a first mission is presented, and (ii) a module where the student is invited to choose an instructor or companion agent from the set of eight different animated agents.

The presentation of the first mission includes illustrations from Istanbul and traditional Turkish music. A male speaker voice tells that there is a chief editor in London who will be his or her companion. The chief editor will formulate the missions, orient the journalist (the student) and provide necessary information at critical stages. The journalist (the student) is to report back to the chief editor who will evaluate the reports and tell what is well done and what needed more work. In the Companion Version the student is, instead, told that there will be a companion journalist with whom s/he will conduct the missions. The student is also told that it is important to cooperate with the companion who, on the one hand, is not completely reliable when it comes to knowledge but, on the other hand, has some of the keys necessary to complete the missions.

5.1.3 Procedure

1. The three experimental leaders introduced themselves to the class as researchers from the university, investigating educational media for the future. Students were told they would be welcome to participate in a study. It was emphasized that the purpose was to listen to students’ opinions on future educational media – full anonymity was ensured. The students were instructed to come, one at a time, to a small room behind the classroom, and in most cases, all students in the class participated.

2. Each participant was randomly assigned to one of four conditions: (i) companion version; realistic intro; (ii) companion version; iconic intro; (iii) instructor version; realistic intro; (iv) instructor version; iconic intro. There were 22-23 participants in each condition. After welcoming a participant and asking what grade s/he was in, s/he was asked to sit down at the computer, press start and follow the instructions.

3. The chosen agent was enlarged and centered on the screen, and the other agent animations disappeared.

4. The participant was asked the open-ended question: “Why did you choose the instructor/companion you did?”

5. Upon answering, an experimental leader reversed the program to bring forth all eight agents again. The participant was then asked whether there was any figure that s/he would definitely not have chosen as instructor/companion. The participant was presented with the eight agents again.

The eight agents were developed out of four basic figures (two male and two female) in 3D Studio Max 5, and their faces were created with the plug-in module FacialStudio. Each of the four basic figures was then rendered in one realistic version (a 3D-figure created with the 3D Studio Max 5’s default renderer) and one iconized version (an illustration rendering created in the FinalToon rendering system). The eight agents had similar skin and hair color. Their facial forms were as similar as possible, given an ambition that they should look like four different individuals. Age related features were held constant. Body shape and eye color on the male agents were identical. The female agents had small differences in body shape; all had middle long hair and the same eye color. Clothing was simple and discrete in all agents. The underlying design rationale was to make the agents as neutral as possible with respect to visual stereotypes (attractiveness stereotypes, personality stereotypes, gender stereotypes, etc.) in order to minimize such influences on the agent choice (see [5]). The animation of the agents was parsimonious and included no sudden movements that would be likely to attract attention. The movement patterns were similar between agents but with a displacement so that movements from different agents would not coincide. Each animation lasted five seconds but was repeated in a loop so that the agents seemed to move continuously. The agents did not speak themselves, but the presentation was accompanied by a speaker voice (introducing the mission) and background music.

Upon completion of the introduction it is time to choose an agent (instructor or companion). In both versions the same eight animated agents, four iconized and four realistic, were simultaneously placed in an oval on the screen (Figure 3). The placement of the agents was randomized before the session and stored in a table accessible to the program.

Figure 3. Example screen dump of the agents used in study 2.

Upon answering, an experimental leader reversed the program to bring forth all eight agents again. The participant was then asked whether there was any figure that s/he would definitely not have chosen as instructor/companion. The participant was presented with the eight agents again.

6 One operated the program and assisted the student; one observed choices and took notes; one conducted the interviews.
then asked what s/he thought to be the important differences, if any, between the figures.

6. Next, one of the experimental leaders presented two scenarios: (i) one with an instructor/companion that focused on the mission and stuck to this (task oriented agent), and (ii) one with an instructor/companion that was more social and, apart from working on the tasks, also supplied information about him or herself in the pauses, telling about former missions, family, friends, interests, and so on (relation oriented agent). Participants were then asked which of the two agents they would prefer, whereupon they were asked to motivate their choice.

7. Finally, two learning style inventories were completed. The total time for a session was on the average eight minutes. After completion, the participant was offered refreshments, was debriefed and thanked for valuable help.

5.2 Results

5.2.1 Measures
For each participant the program logged: the condition (i) - (iv) (see section 5); the positions of the agents; the time it took for the participant to choose an agent; the chosen agent and its position. The qualitative data, that is the participant’s articulations and motivations of their choices, were noted manually during the session and transcribed within a few hours. In coding the answers to the question: “Why did you chose the agent you did?”, all three coders independently chose the following categories for the arguments: aesthetic, personality, gender, and other/none. Some answers were classified in two of the categories. Upon comparing the classifications made by the three coders, a few differences occurred. After discussion, a joint result was arrived at (Table 2).

Table 2. Categorization and distribution of arguments for choice of agent (translated from Swedish).

<table>
<thead>
<tr>
<th>Arguments</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aesthetics</td>
<td>24</td>
</tr>
<tr>
<td>Personality</td>
<td>29</td>
</tr>
<tr>
<td>Gender</td>
<td>5</td>
</tr>
<tr>
<td>Aesthetics + Personality</td>
<td>6</td>
</tr>
<tr>
<td>Aesthetics + Gender</td>
<td>1</td>
</tr>
<tr>
<td>Personality + Gender</td>
<td>5</td>
</tr>
<tr>
<td>Other / None</td>
<td>20</td>
</tr>
</tbody>
</table>

5.2.2 Visual Style Preferences
Regarding visual style preferences, 63% of the participants (69% of the females, 57% of the males) chose an iconized agent, and 37% of the participants (31% of the females, 43% of the males) chose a realistic agent.

Examples of the participants’ motivations for choice of agent can be seen in Table 3 below.

Table 3. Categorization (cf. Table 2) and number of arguments for motivating choice of agent.

<table>
<thead>
<tr>
<th>Category (No.)</th>
<th>Arguments (typical examples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aesthetics (24)</td>
<td>“I think the other ones look a bit strange.,” “[…] because it is a 3D-figure, that has much more style than a drawn one.,” “I didn’t want a 3D, I think they look a bit strange. The 2D ones look better, more trendy.”</td>
</tr>
<tr>
<td>Personality (29)</td>
<td>“He looks as if he has good self-confidence. Some others look insecure and shy.,” “I thought he looked kind.,” “He looks like someone who could teach you something.”,” “She looked kind and reliable – not someone that will shout at you.”,” “It looked most sympathetic.”</td>
</tr>
<tr>
<td>Gender (5)</td>
<td>“I wanted it to be a woman, because there are too few woman bosses. As long as it is a woman it is not so important which one of the figures.”,” “I don’t know, but I am a girl and want to work with girls.”,” “[…] because he is male. Maybe there is criminality where one is going, and with a guy the risk of being attacked is less than if you are a boy and a girl.”</td>
</tr>
<tr>
<td>Aesthetics &amp; Personality (6)</td>
<td>“She looks good graphically, and she also looks like a chief editor.”,” ”She was the best looking one, and also kind of fair.”</td>
</tr>
<tr>
<td>Aesthetics &amp; Gender (1)</td>
<td>“[…] because it is a girl and I think they look better. They are more ugly when they are computer made.”</td>
</tr>
<tr>
<td>Personality &amp; Gender (5)</td>
<td>“[…] because it is a girl, and because she looks happy and very kind.”,” “[…] I chose here because I think it is fun with a girl if you are going to have a good cooperation, and she looked nice and the most human.”</td>
</tr>
<tr>
<td>Other (5)</td>
<td>“It looked most normal and common.”,” “[…] because it is a computer game and I like computer games. I chose the one I think is best animated by the computer.”</td>
</tr>
<tr>
<td>No reasons (15)</td>
<td>“It doesn’t really matter – some of them look a bit special, but honestly – who cares?,” “I don’t think it matters who is boss as long as he or she is kind, and you cannot see that from the outside.”</td>
</tr>
</tbody>
</table>

As a part of the examination of a possible relation between a preference for iconized agents and a view of agents as social actors, we focused at the categories “aesthetics arguments” vs. “personality arguments” and related them to the choice of visual style (see Table 4, next page). A strong relationship was found between personality arguments and a preference for iconized agents.

Workshop Proc: *Adapting the Interaction Style to Affective Factors*, 10th International Conference on User Modelling (UM’05), Edinburgh, Scotland, July 24-29, 2005
5.2.3 Social Style Preferences

Looking at social style preferences, 59% of the participants (65% of the females, 52% of the males) chose a relation oriented agent, and 41% of the participants (35% of the females, 48% of the males) chose a task oriented agent.

As to the arguments for choice of agent, 72 answers were considered after sorting out motivations such as: “That just seems better.”, “I just think so.”, and a few responses of silence.

Seventeen participants motivated their choice of a (task and) relation oriented agent as it being: “[…] more fun, nice or interesting.”. Another eighteen participants gave this argument in a more elaborated form – some of them explicitly spoke about the importance of “[…] personal relations.”. Eight participants motivated their preference for a relation oriented agent as: “[…] being more playful and easy-going – the other (task oriented) agent would make the task too serious and hard.”. Three participants, finally, motivated their preference for a relation oriented agent in terms of what is “[…] normal or common.”. Two of those found the relation oriented agent more interesting because it is: “[…] not tiresome and a nuisance.”; another seven pointed at the risk of getting distracted; the third group of seven participants held that a relation oriented agent would be “[…] trying, tiresome and a nuisance.”; another seven pointed at the risk of “[…] getting distracted.”; the third group of seven participants spoke of a relation oriented agent as one that “[…] does unnecessary or meaningless things instead of focusing on what is important.”. Finally, five participants stated that the agent “[…] is a computer character and not a human being, and that they therefore didn’t want to know any personal things about it.”.

Two participants explicitly answered that the ideal would be to have both versions available to choose from: “Sometimes you feel like talking, but sometimes you prefer a companion that is quiet and sticks to the task – the best would be if one could choose between companions that have different personalities.”; “It depends – sometimes I would like one that is talkative and social, but sometimes I cannot stand that.”.

5.2.4 Visual and Social Style Preferences Together

Participants who preferred a relation oriented agent tended to prefer an iconic style agent ($\chi^2(1) = 13.75; p = 0.00$), whereas participants who preferred a task oriented agent did not reveal any significant preference ($\chi^2(1) = 0.24; p = 0.62$), (Figure 4).

6. CONCLUSION AND DISCUSSION

A central question in the two studies was whether any relationships could be found between a) users’ choice of visual style in agents with respect to realism versus iconicity and b) users’ choice of social style with respect to task orientation versus (task and) relation orientation. In particular, the hypothesis that iconized visualization is better suited for representing a relation oriented, more subjective, agent – and therefore preferred by users who prefer a relation oriented agent – was tested.

Together, the results of the two studies provide some support to the hypothesis, even if further investigations are required in order to secure as well as to better understand the results.

A relationship appears in the studies between a preference for an iconized agent and a preference for a relation oriented agent. If, as we propose, this result is interpreted in terms of iconized agents being more easily conceived of as subjective, socially rich characters than realistic agents – a next step is to ask what the underlying cognitive processes in that case could be. A possible starting point is Laurel’s [7] comparison of computer characters and theatre characters, declaring it as central in both cases that the characters function as stereotypic “shorthand” for understanding and predicting behavior, rather than as full-blown personalities. And “the artistic side of the design problem” means, according to Laurel, to represent the agent to the user in such a way, that appearance is shaped to suggest the internal traits of the dramatic character or the agent. Elaborating on this, an iconized character in contrast to a realistic one prompts the user to develop the character. A highly realistic agent is a visual – and social – fact, which does not leave much for a user to imagine. It is like an objective statement, whereas an iconic agent can be elaborated by the user, who may fill in and create from his or her own subjective experiences. Thus, someone who is interested in understanding and creating psychological and personal issues and prefers a subjective, relation oriented agent, may also prefer an iconic be-

---

7 Actually, in the case of theatre, the visual shaping of characters to a large extent makes use of visual stereotypes and heavy make-up – and the large distance to the scene reinforces the appearance of the artists on the theatre scene as visually iconized.
fore a realistic visual style, if offered a choice. This line of reasoning can also be associated to McCloud’s thesis [10] that it is easier to identify with an iconic than a realistic character. A realistic agent is more of an object, a finished existence. An iconic agent is more subjective, enabling a user to float into the agent with his or her own experiences.

Furthermore, the series of studies by Reeves and Nass [12] show that a computer with only minimal visual or auditative cues – such as a pictured mouth and eyes or a voice – triggers social projection in the sense that humans treat computers as social actors. The present line of reasoning could be said to extend the conclusion from Reeves and Nass [12] one step, stating that social projection is not only likely with a simple visual representation, but sometimes even more likely than with a complex visual representation.

Thus, a tentative conclusion with respect to design is that if the goal is to design a pedagogical agent, rich in subjectivity and more relationally oriented, iconic visualization may be the better choice. Likewise, if the goal is to design an objective, task oriented pedagogical agent, a realistic representation may fit better.

Another outcome of the studies regards user variability. They show a considerable variability in user preferences. As regards social style (relation oriented – task oriented) this can be related to Bickmore’s [2] observations that REA in the social condition evoked strong and diverging reactions. Several subjects “reported liking the social dialogue aspects of the interaction: […] It wasn’t just real estate talk, so I felt like it made her more human […] It sounds like she’s on your side when she says things are expensive” ([2], p. 84). Other subjects didn’t like it at all: “I come in and I shop and I get the hell out. She seemed to want to start a basis for understanding each other.” ([2], p. 85).

Also, studies of Laura [2] indicate that user appreciation of her relation orientation ranged widely. According to Bickmore [2], a representative positive comment was the following: “I like talking to Laura, especially those little conversations about school, weather, interests, etc. She’s very caring. Toward the end, I found myself looking forward to these fresh chats that pop up every now and then.” They make Laura so much more like a real person” ([2], p. 184.) A representative comment from someone who was clearly of another opinion was: “I didn’t really like Laura very much […] Actually I liked all of the software except for the animated conversation thing.” ([2], p. 185).

Furthermore, Bickmore and Cassell [3], attempted to relate the varying user reactions towards REA to user characteristics, more specifically to the introversion-extroversion dimension. A significant difference showed up, both in the case with an embodied REA, and when there was only phone conversation, in that “over-all extraverts liked REA more when she used social dialogue, while introverts liked her more when she only talked about the task” ([3], p. 21, emphasis added). In our second study, user characteristic data regarding learning styles have been collected. The analysis of this data will appear in coming articles.

Finally, given the apparent variability in user preferences and reactions, how can the issue of adaptation of the interaction be approached? Given the current state, it is hard to envision a system that in itself would adapt to a suitable visual and/or social style. So, for the time being, one simple “adaptation” is to present users with alternatives and let them choose – if for no other reason than the mood they happen to be in from one day to another.

7. ACKNOWLEDGMENTS
This work was supported by the Swedish Research Council, VR.

8. REFERENCES
PAPER III:

AESTHETIC STEREOTYPES AND VIRTUAL PEDAGOGICAL AGENTS

ABSTRACT

The paper deals with the use of visual stereotypes in virtual pedagogical agents and its potential impact in digital electronic environments. After an analysis of the concept of visual stereotypes, affordances as well as drawbacks of their use in the context of traditional media are analyzed. Next, the paper explores whether virtual pedagogical characters introduce anything novel with respect to the use of visual stereotypes – as compared both to real life interaction between humans and to the use of visual stereotypes in traditional non-interactive media such as magazines, film, TV and video. The answer is that novel affordances, as well as novel drawbacks, indeed are being introduced with the use of visual stereotypes in virtual characters. Finally, we suggest some future directions of research, and from an educational and societal perspective we argue for the imperativeness of such research.

Keywords:
Pedagogical agent, Virtual character, Visual stereotype, Learning, Gender

Introduction

Virtual pedagogical agents, i.e. computer generated characters in pedagogical roles are entering the digital society in increasing numbers. They are found in educational programs, from preschool to university. They are also found in broader educational contexts as virtual medical counsellors, physical exercise coaches and guides on city homepages, and they appear in edutainment and infotainment settings.

In recent years there has been a growing focus on the social dimension of interaction with virtual agents, with a variety of social competences being simulated in various agents (Baylor et al., 2005; Bickmore & Cassell, 2005; Gulz, 2005; Hall et al., 2004; Johnson, 2003; Paiva et al., 2004). Behind this lies the influential framework of “computers as social actors” by Reeves and Nass (1996), based on a large number of studies that show that people spontaneously apply social interpretations and conventions when interacting with computer based media. This human disposition, it is argued by Reeves and Nass and many others, ought to be further exploited in computer system design in order to make interaction smoother and more satisfying.

In human-human social interaction the visual appearance of other people is known to play a central role, with profound effects on our attitudes as well as behaviour. On the one hand, there are dynamic visual aspects such as gestures, facial expressions and gaze, which are extensively researched within the agent community. On the other hand, there are static visual aspects such as body and face properties, skin, hair and haircut, clothes and attributes. In spite of their documented impact in human-human social interaction (Kalick,1988; McArthur, 1982), the static visual aspects – that we will focus on in this article – have been very little attended to in research on virtual agents (Gulz & Haake, 2006a; Gulz & Haake, 2006b).

Submitted to *Journal of Educational Technology & Society* (June, 2006)
Particularly in interacting socially with unfamiliar others, humans exploit visual cues – static as well as dynamic – to form expectations for guiding the interaction. In drama theory the concept *physical personality* refers to the aspects of a drama character’s appearance that immediately produce an impression of personality and initiate a set of expectations and attitudes. Among those are many static visual aspects such as body shape, height, sex, race, face, hair, clothing, make-up and facial hair (Brahnam, 2001). In contrast to how these cues are immediately picked up in the encounter with another human being, the linear stream of spoken information is incredibly slow (even though, of course, some voice characteristics are also quickly picked up). 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”.

These categorization processes, in which we quickly form expectations on a person’s likely behaviour, attitudes, opinions, personality, manners, etc., rely heavily upon *stereotypes* – that is, ideas of “typical” representatives of certain human categories, where clusters of properties are ascribed to these categories and their representatives. *Stereotype* is one of the main concepts in this article, and focusing on the visual static aspects of stereotypes, we will use the term *visual stereotypes*.

The objective of the article is to explore the use of visual stereotypes in virtual pedagogical agents. Firstly, we ask to what extent affordances and drawbacks found in the use of visual stereotypes in real life interaction and in traditional media, reappear when used in virtual agents. In attempting to answer this question, we also present and discuss some of the (few) empirical studies carried out in the area. Secondly, we analyse in what ways interaction with virtual pedagogical characters introduces something *novel* with respect to the use of visual stereotypes, both in terms of affordances and of drawbacks. But before entering these two issues, we provide a background by analysing the concept of *visual stereotype*, relating it to that of *visual prototype*, and discussing visual stereotyping with respect to *visual naturalism versus visual stylisation*.

**The concept of visual stereotype**

![Figure 1: Examples of visual stereotypes.](image)

Figure 1 presents four examples of *visual stereotypes*. Many observers will in these pictures see a teenager (1a); a housewife (1b), a craftsman (1c); and an air-hostess (1d). A visual stereotype, in our sense of the term, consists of a number of visual attributes in a person that will make a majority of observers perceive the person as an *illustration*, or a *typical instance*, of a human group: a professional group, a social group, etc. That is, the *visual input* activates expectations on other – not visible – attributes in the person: how he/she is likely to behave and to talk, what he/she can be expected to say or not say; what attitudes and opinions he/she will be likely to have, etc. In this way visual cues carry social baggage.

Submitted to *Journal of Educational Technology & Society* (June, 2006)
Gender is often an important aspect of visual stereotypes. A “typical craftsman” is a man, whereas a “typical air hostess” is a woman (see figure 1). Furthermore, a representation of a “typical scientist” is at the same time also a representation of a “typical male scientist”, whereas a “typical female scientist” is a stereotype of its own. Correspondingly, the visual stereotype of a “sweetie” concords with that of a “female sweetie”, whereas “male sweetie” is another and separate visual stereotype. Gender will be a recurring issue in this text.

Stereotypes and prototypes

Within the cognitive sciences the concept prototype is used to stand for “a typical exemplar” of a concept. For example, an apple is a prototypical fruit whereas a kiwi is not, and a lawyer is a prototypical law person whereas an investigation secretary is not. Focusing on the visual aspects, one may also speak of visual prototypes in parallel to visual stereotypes.

In the context of this article, an interesting difference between the concepts is that a (visual) prototype is something neutral, whereas a (visual) stereotype brings along negative associations (Schneider, 2003). A picture or description of a prototypical youngster, a prototypical nurse or a prototypical beauty is not necessarily something negative. The presumed basis for a certain exemplar being prototypical is the frequency of occurrences in human perceptions and experiences, and in that sense a prototype relates to something “that is really there”. A stereotype, on the other hand, is often associated with something that culture, media, etc. has constructed out of dubious starting points and where there is nothing “that is really there”. (A complicating factor, that blurs the distinction just given, is of course that media constitutes an important part of our perceptions and experiences).

We will in the following use the term visual stereotype but attempt to assign to it some of the neutrality from the term prototype. Our purpose is, namely, to study the use of visual stereotypes from two angles and look for positive affordances as well as drawbacks. Thus, we are not assuming that the use of visual stereotypes is necessarily evil, unjustified or to be combated. (Actually, it was not until the 19th century that the word stereotype – originating from the Greek words stereos for “solid” and typos for “a model” – became linked to prejudice and discrimination (cf. Schneider, 2003).)

As to the bases of visual stereotypes, our standpoint is that it is a heterogeneous phenomenon. For some visual stereotypes the expression “no smoke without fire” applies. Culture and media may reinforce and exaggerate them, but there is “something there” behind the visual stereotypes. For instance, a punk look indeed often goes along with an individualistic attitude and a desire to be allowed to go ones own way and not be forced to follow established societal norms. In the case of other visual stereotypes, the world has changed substantially compared to the situation from which they originate, and there is no or very little accuracy in them today: e.g. the personality and behavioural habits of the eccentric British explorer in his Kaki shorts, short-sleeved shirt, pith helmet, and brown shoes. Yet other visual stereotypes were from the start pure constructions. The “stupid blonde” stereotype is one such example. It certainly is not, nor has ever been, the case that blonde women on the average are more stupid than non-blonde women are.

Stylised visual stereotypes

Visual stereotypes can appear in many different media formats: photos, movies, paintings, drawings, comics, animated movies, etc. Some of these formats allow different degrees of visual naturalism, which in the context of stereotypes is a feature of interest. There is a whole scale from photorealism on the one hand to pronounced stylisation on the other, e.g. in a cartoonish style or other artistic style that modifies and often simplifies a persons’ appearance (Gulz & Haake, 2006b).
Stylisation makes it possible to sharpen and exaggerate a visual stereotype (see Figure 2). Such amplified visual stereotyping via stylisation is foremost associated with graphical media, such as cartoons and animated motives. But it is also used in theatre, where the shaping of characters sometimes makes extensive use of visual stereotypes. Heavy make-up and large distance to the scene can, furthermore, reinforce the appearance of the artists on the scene as visually stereotyped. In classical Chinese opera and in Comedia del Arte, for example, dresses and make-up as well as gestures are pronouncedly stylised (see Figure 3). Also, Hollywood dramas can rely upon stylised visual stereotyping. As an example, Indiana Jones (Figure 3) wears throughout all three movies more or less the same outfit, characterized by his fedora, leather jacket, unbuttoned shirt, and bullwhip, which function as immediate cues for identification. Likewise, it is often easy in Hollywood dramas to predict the roles of characters and of who is going to die or survive on the basis of various visual and exaggerated characteristics.

Visual stereotypes in traditional media characters as well as in virtual characters

Visual stereotypes – affordances

Visual stereotypes are an important aspect of human thinking in their function as cognitive short cuts for making action and life tractable for human beings. Instead of becoming overwhelmingly occupied with thoughts and questions about people that we encounter, we make use of their visual appearance to situate them, in order to focus on interaction as such (Brewer, 1988). That is, visual stereotypes frame our expectations. They are also used for building common references in conversations about other people. In brief, they are a navigation tool in a social environment that would otherwise be overwhelmingly complex and demand a practically insurmountable burden of processing (Smith & Medin, 1981).

In traditional visual media – theatre, film, comics – it is essential that readers or spectators are scaffolded to gain an idea of the characters (their personalities, habits, manners, opinions, predispositions etc.).
Without such starting points for entering the story, many plots would simply not be possible to follow. And here, as Laurel (1993) points out in discussing theatre, the visual appearance of characters can be used to “suggest the internal traits of the character in order to function as shorthand for understanding and predicting the character”. As mentioned above, some forms of theatre indeed drive this very far.

Consequently, a parallel use of visual stereotypes in digital media may provide starting points both for the interpretation of virtual characters and for the interaction with them. Adequate starting points unleash resources to focus on the content of the interaction, afford smoother interaction and may generate a greater sense of enjoyment and accomplishment in users.

This line of thought is indeed reflected in the many design recommendations or guidelines for virtual characters (as well as for traditional media characters), that underline the importance of consistency between features such as voice, gender, looks and role of a character (Nass et al., 2000). With a visual appearance that corresponds to behaviour and personality predictions, user get their expectations acknowledged and leveraged, and the interaction becomes smooth and efficient.

However, choosing an adequate visual stereotype is not always easy, as exemplified by the following case provided by De Rosis et al. (2004). A virtual character was to be designed 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, the designers also became aware that the appearance and behaviour of the character disturbed these users, and designed a new character with a more professional communication style and more classical attire.

The point is that the first visual stereotype, the young attractive female secretary character, was not an adequate starting point for the users of the system – but disturbed and distracted. Furthermore, visual redesign was indeed required. It would not have been sufficient to redesign the dialogue and gestures but leave the visual appearance intact.

Another example involves two different design choices of a virtual assistant for a city home page. In Botkyrka, a Swedish city with a high percentage of immigrants, a stereotypically “Swedish-looking” light-blond female character was introduced as virtual assistant on the site. However, due to negative user responses the character was removed. In contrast, in the Swedish city Malmö, equally a city with a high percentage of immigrants, the character Sara (figure 4) was chosen. In this character stereotypical ethnicity attributes, as well as stereotypical gender attributes, were treated carefully. Notably, the discussions behind the visual design of the Sara character were extensive.

![Figure 4: The virtual city guide Sara.](image)

Visual stereotypes – drawbacks

A visual stereotype can – by its nature – activate misleading expectations. Even if based on frequency distributions of property-clusters in peoples’ experiences and thus corresponding to “actuality” in a statistical sense, a visual stereotype can in a given instance be inadequate and misleading. In real life an example would be a youngster whose street fashion look signals “tough, rebellious and cheeky” but who actually is very shy and timid.
Correspondingly, an unfitting visual stereotype used for a film character or a comic character – when not an intentional choice by the producers – may confuse and irritate users and induce an impression of a non-believable and un-professionally staged character. Consider all the fuzz around the choice of James Bond actors, where some last for only one production. Others reappear again and again and become more or less synonym to the role, especially Sean Connery (alternatively Roger Moore) who almost equals James Bond. Even though there are other factors involved, the visual appearances play a large role in these outcomes.

For the case of interactive media consider the Botkyrka example reported above, where the chosen visual stereotype of a “very Swedish-looking” women, activated undesired and misleading expectations as to the aim and use of the city home page.

Another important drawback is that visual stereotypes may be perceived to represent the normal, and make visual appearances that diverge from the stereotype be perceived as odd, unusual, or even abnormal. For instance, a somewhat thin and spectacled craftsman is “no real craftsman”. In this way, visual stereotypes can hide or suppress nuances and an existing manifold. There is also an aspect of self-reproduction and self-reinforcing of stereotypes due to the close interactions between media and “real life”, which can have unfortunate effects. Societally undesirable gender stereotypes of a normative kind are frequent in traditional non-interactive media, and can be observed in the case of virtual characters as well. For instance, many computer game characters reproduce visual stereotypes. In 1998 it was concluded in the Next Generation magazine that despite dramatic increases in the number of female game characters, “they all seem to be constructed around very simple aesthetic stereotypes. In the East, it’s all giggling schoolgirls and sailor uniforms, but in the West the recipe appears to be bee-sting lips, a micro-thin waist, and voluminous, pneumatic breasts” (p. 8). And even though there has been some change, overall there is still truth in this analysis. From more recent discussion forums one can learn that some female gamers refuse to play female characters and feel insulted by how they are designed, and also that some male gamers are unsatisfied with the masculine stereotypes presented (visually and otherwise): ‘I usually play as female characters, because male characters are always hyper-masculine and that’s not how I feel. However, if there’s a feminine guy, I WILL choose him’ (GameGirlAdvance, 2004).

In sum, we have discussed a number of advantages and drawbacks with use of visual stereotypes, where what is known from traditional media reappears in the virtual world. We now proceed towards what comes as novel with virtual characters.

Novel affordances and risks with virtual characters and visual stereotypes

Introducing novel possibilities

All sets of visual cues that can appear in real human beings, or in photos and films of real humans, can also be made to appear in a virtual agent. Thus, all visual stereotypes that can be seen in live human beings or in traditional media portraying humans can also be reproduced in virtual characters. But there are additional possibilities in virtual characters due to the extended degrees of freedom regarding visual shaping. In virtual characters it is easy to “cut and mix” and arrive at combinations that do not occur, or rarely so, in real human beings. Thus, it is relatively easy to challenge, or break down, visual stereotypes: to combine visual elements from different stereotypes or to combine a given visual stereotype with an unusual role.

It can be argued that such playing around with visual stereotypes is just as possible in other graphical media, such as comics and animated movies. Nevertheless it seems, that in practice interactive media has brought out this more extensively. While traditional graphical media relies on the observer or reader as a passive consumer of pre-designed stereotypes, the interactive virtual arena activates the participants. In the area of computer gaming, we find communities where players themselves contribute to the design and development of characters. Here a remarkable character diversity can be observed. As to gender, several new appearances of female heroines, androgynous characters and other kinds of in-betweens, have come into existence (Schleiner, 2000).
This points towards the potential of using virtual pedagogical characters as visual stereotype busters, to present the non-standardized and expose a manifold in combinations of ethnicity, social classes, gender, professional roles, and so on. Such offering of a broader range of styles and identities may, furthermore, enable social identification and role modelling – something known to be of importance in learning contexts – for a larger number of students. Figure 5 shows some avatars (Sims) from the virtual world Second Life, out of which some are used to explore alternative gender and personality.

**Dilemmas with novel possibilities**

The idea of breaking with visual stereotypes for pedagogical purposes can, however, be in conflict with the pedagogical exploiting of visual stereotypes to facilitate smooth and efficient interaction. Because, as observed earlier in the text, the use of visual stereotypes may, by acknowledging and leveraging learners’ expectations, enable learners to interact smoothly with a character and focus on the learning activities and materials in question – rather than making learners confused and distracted by unexpected features and behaviour in a character.

The goal of “smooth and efficient communication” is central in the virtual agent research domain. In light of this it is intelligible that Moreno et al. (2002) highlight the aim of obtaining “pedagogically effective animated agent” and the issue of “how stereotypic information [in the sense of visual stereotypes] can be used in order to facilitate learning from animated agents”. Nevertheless, there is a questionable lack of problemizing both the issue and the results of their study (2002), which indicate that participants learn significantly more from the male virtual tutors on the subject of blood pressure than from the female virtual tutors. The proposed explanation for the outcome goes that “the female tutors broke with rules of etiquette about who should teach at a college level by not conforming to the stereotype of males as professors”, and is left without further comment. That is, there is no mentioning of a conflict between, on

Submitted to *Journal of Educational Technology & Society* (June, 2006)
the one hand, a wish to exploit the male professor stereotype in order to “facilitate learning from animated agents” and, on the other hand, a wish not to further strengthen the idea about the male professor as the norm by using this stereotype. Likewise, the authors, without comments, pose the question: “Do people learn more effectively about car repair form an agent named Joe who wears greasy overalls, or can they learn just as effectively about this topic from an agent named Nancy in a pink apron?” . Compare figure 6 showing an extract from the Joe Doe instruction series for US Army, made during the Second World War by Will Eisner.

In contrast to the Moreno et al. study (2002), the complexity and hidden dilemmas in the use of visual stereotypes come forth clearly in the work of Baylor and her group. Baylor directs the project “Challenging Stereotypes toward Engineering with Pedagogical Agents” within the Program for Gender Equity in Science and Engineering. The group investigates, among other things, the influence of character appearance on female choice of engineering subjects. In one study (Baylor & Plant, 2005), the researchers compared effects of two different female engineer characters. One character was designed as a stereotypical female engineer (based on pre-test validation): “geeky”, greyish and homely. The other character, Nina, was designed as attractive, feminine and outgoing. In other words, Nina countered the expectation that an engineer, whether male or female, is not (prototypically) feminine. Results were that female students who worked with Nina were more interested in the subject than those interacting with the ‘stereotypical’ character and also more likely to believe that they could themselves be successful as an engineer. The result is, however, problemized in Baylor (2005), where it is pointed out that “designing interface agents in non-stereotypical roles (e.g., the “sexy” engineer agent, Nina) may lead to desired attitudinal effects, but could have more detrimental effects in the long-term for females’ beliefs toward women in demanding careers”. Another study (Baylor, 2005) involved a female and a male engineer character in the role of coach for an engineering practice tutorial. Here, attitudinal and learning outcomes were superior for students who worked with the female coach, in other words the coach that broke with the (visual) stereotypes of an engineer. What, according to the author, seemed to happen was, however, that the prejudice of females as less competent in technical domains spills over to the virtual area, generating increased self-efficacy of the kind “If she is able to do it, I can do it!”. And these prejudices about less competent female engineers are, Baylor (2005) remarks, not ones that one would like to reinforce and disseminate.

In sum, there is a need to handle dilemmas in which, on the one hand, the use of a visual stereotype can contribute to efficient communication in a pedagogical situation by leveraging users’ expectations but where, on the other hand, the breaking with the stereotype can be desirable from a societal and a longer term pedagogical perspective. Likewise, there is a need to handle dilemmas where the breaking of a visual stereotype may have positive attitudinal and learning effects, but at the same time produce or reinforce questionable conceptions, as in the “sexy” engineer example above.

**Introducing novel risks**

As to the detrimental normative function of visual stereotypes, as discussed in a previous section, the extended degrees of design freedom offered by the virtual world brings about additional risks. The construction and promotion of idealized super people with “perfect” bodies and looks (and even lives) has since long been abounding in non-interactive media such as TV, video and magazines. Figure 7 shows the ‘femme fatal’ P’Gell in Will Eisner’s comic Spirit (1940-52). In this sense there is not much new under the sun, when interactive computer media continues this portraying of the ideal by promoting stereotypic instances never found in real life, no matter how much cosmetic surgery, such as the big-breasted, thin-waisted action heroine Lara Croft: a biological contradiction whose tiny abdomen could never house all her vital organs. In particular not if she also is to perform spectacular stunts and engage in violent fighting, instead of fainting like the corseted females of the 19th century (figure 7).
Nevertheless – this portrayal of the ideal can be taken one step further with interactive computer media. A key difference lies in what is otherwise seen as a central potential of virtual characters – not the least in pedagogical terms – namely their interactivity: virtual characters may communicate, respond, and answer. A possible effect of this is that the distance between users (learners) and those “ideal super people” is diminished. Until now we have watched, and read about, fabulous, good-looking people in movies and magazines (cf. P’Gell in figure 7). If we are also to interact with them – in an era already desperately pursuing perfection in appearance – this might have detrimental effects on peoples self image and self esteem, as the interactivity may blur the distinction between “artefact” and “reality”.

The role of visual naturalism versus stylization is, furthermore, interesting in this context. With stylization it is, as pointed out, possible to sharpen and exaggerate visual stereotypes, including perfection stereotypes. On the other hand, the standpoint “this is another kind of creature that I do not have to live up to and compete with” intuitively seems easier to maintain in the case of stylized rather than naturalistic representation. However, comics specialist McCloud (1993) holds that identification with a cartoonish character comes much easier than identification with a realistic character. Also, Reeves & Nass (1996) argue that humans do not “keep a distance” even to the most ersatz representations of a person. In other words, the question on how visual naturalism/realism vs. stylisation affects distance and identification is an open question.

Another, related risk concerns the fact that virtual worlds involve users in a more active way than movies, books, etc. You can participate in various activities, including many simulated everyday activities – c.f. Second Life or Entropia Universe – and furthermore these activities may go on and on. There is no ending, as in the movie or book. This increases the risk (or potential) for users to indeed enter into “another world” with its characters and in absorption, leave reality behind. The addiction risk is apparent and extensive and “absorbing” interaction with stereotypical characters may have negative consequences for people’s conceptions of real people and real social life.

**Conclusion**

Positive and negative sides in the use of visual stereotypes known from real life interaction and from traditional visual media, reappear in interactive media. There is, for instance, their invaluable function as cognitive tools for handling a complex social environment, and there is their problematic normative function that can make what diverges from a visual stereotype be perceived as odd or abnormal.

But there are also novel possibilities and risks introduced by visual stereotypes in virtual agents. We have emphasized the possibility to break with, or exploit, stereotypes for pedagogical purposes; not the least...
for enabling exploration of a broad range of identities and extend possibilities for social identification and role modelling. As to risks, we have pointed at the replenished risks that follow from users’ interacting with, and perhaps being absorbed with, visually idealized stereotypes.

In view of the pedagogical and societal relevance of the phenomena described - the potentials as well as risks - further research is required on:
- identification with virtual characters, and role modelling effects
- possible pedagogical effects by challenging visual stereotypes
- how perfection in appearance in interactive virtual characters can affect users (learners).

For each of the issues increased knowledge on visual details, such as for instance the naturalism-stylisation aspect, is required as well.

Indeed, there is interesting research in the case of traditional, non-interactive, media relating to the issues listed above. But given the additional affordances – in positive and negative senses – of interactive media, the questions must be separately approached for interactive media. Reeves & Nass (1996) hold that traditional media, such as television, and new media, such as computers “afford the same problems and opportunities of stereotyping” (p. 170). We think that there are many similarities – but also differences. Not the least, negative effects from media and visual stereotypes can be reinforced with computer media. Simultaneously, if academic research in the domain keeps up with the technological and commercial development, there is a potential to take more responsibility and be proactive in canalising some of the development.

**Acknowledgements**

The authors are grateful for the financial support of the Swedish Research Council, VR. Also sincere thanks to Daniel Labbé for many valuable comments on the article.

**References**


Submitted to *Journal of Educational Technology & Society* (June, 2006)


Submitted to Journal of Educational Technology & Society (June, 2006)